

Genature version 5.1.3
Copyright (c) 1993-2002 CompuGen Ltd.

OM protein: protein search using sw model

Run on: November 6, 2002, 14:04:59 : Search time 10.451 Seconds
(without alignments)
312,880 million cell updates/sec

Title: US 09 834-409-4

Perfect score: 751

Sequence: 1 G1SGSPPTTINRISVST ANNMWPTPLPTGVSEFLE 134

Scoring table: BLOSUM62 Gapop 10.0 / Gapext 0.5

Searched: 246,08 seqs, 244,064 residues

Total number of hits satisfying chosen parameters: 231628

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post processing: Minimum Match 0%

Maximum Match 100%

Listing first 100 summaries

Database: Issued patents AA*

- 1: /seq26/ptdata/1/iaa/5A_0MBR pep*
- 2: /seq26/ptdata/1/iaa/5A_0MBR pep*
- 3: /seq26/ptdata/1/iaa/6A_0MBR pep*
- 4: /seq26/ptdata/1/iaa/6A_0MBR pep*
- 5: /seq26/ptdata/1/iaa/6A_0MBR pep*
- 6: /seq26/ptdata/1/iaa/6A_0MBR pep*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

				SUMMARIES	
Result No.	Score	Query Match	length DB ID	Description	
				Patent No.	Sequence
1	219	29.2	1466	6 5256642 6	Patent No. 5256642
2	219	29.2	1466	6 5472939 6	Patent No. 5472939
3	219	29.2	1537	6 5256642 5	Patent No. 5256642
4	219	29.2	1537	6 5472939 5	Patent No. 5472939
5	219	29.2	1847	6 5256642 10	Patent No. 5256642
6	219	29.2	1847	6 5472939 10	Patent No. 5472939
7	219	29.2	2039	6 5256642 2	Patent No. 5256642
8	219	29.2	2039	6 5472939 2	Patent No. 5472939
9	214	28.5	264	1 US 07-905 983 2	Sequence 2, Appl
10	198	26.4	133	2 US 08-356 461 31	Sequence 31, Appl
11	198	26.4	133	2 US 08-769 967A 31	Sequence 31, Appl
12	198	26.4	253	2 US 08-367 361 29	Sequence 29, Appl
13	198	26.4	254	2 US 08-769 967A 29	Sequence 29, Appl
14	195	26.0	254	2 US 08-356 461 30	Sequence 30, Appl
15	195	26.0	254	2 US 08-769 967A 30	Sequence 30, Appl
16	190	25.3	464	4 US 08-981 234B 2	Sequence 2, Appl
17	160	21.3	484	4 US 08-139 195 2	Sequence 2, Appl
18	159	21.2	169	1 US 08-310 416A 18	Sequence 18, Appl
19	159	21.2	169	2 US 08-888 171 18	Sequence 18, Appl
20	159	21.2	254	1 US 08 410 416A 13	Sequence 13, Appl
21	159	21.2	254	2 US 08 888 171 13	Sequence 13, Appl
22	159	21.2	254	2 US 08 435 149 1	Sequence 1, Appl
23	159	21.2	293	1 US 08 310 416A 16	Sequence 16, Appl
24	159	21.2	293	2 US 08 888 171 16	Sequence 16, Appl
25	159	21.2	470	2 US 08 528 057 42	Sequence 42, Appl
26	159	21.2	474	2 US 08 528 057 44	Sequence 44, Appl
27	159	21.2	477	2 US 08 528 057 2	Sequence 2, Appl

Patent No. 5514787 2	484	21.2	159	23
Sequence 4, Appl	573	21.2	159	25
Sequence 42, Appl	611	21.2	159	30
Sequence 2, Appl	323	21.1	158.5	31
Sequence 14, Appl	324	21.1	158.5	32
Sequence 14, Appl	324	21.1	158.5	33
Sequence 46, Appl	324	20.6	155	34
Sequence 27, Appl	353	20.4	153.5	35
Sequence 27, Appl	197	20.4	153.5	36
Sequence 42, Appl	198	19.6	147	37
Sequence 42, Appl	229	19.6	147	38
Sequence 31, Appl	290	19.2	147	39
Sequence 29, Appl	126	19.2	144	40
Patent No. 5514582	1019	19.2	144	41
Sequence 4, Appl	1019	19.2	144	42
Sequence 4, Appl	1019	19.2	144	43
Sequence 4, Appl	1083	19.2	144	44
Sequence 2, Appl	1083	19.2	144	45
Sequence 2, Appl	1083	19.2	144	46
Sequence 2, Appl	1083	19.2	144	47
Sequence 58, Appl	274	18.6	140	48
Sequence 58, Appl	274	18.6	140	49
Patent No. 5514582	46	18.5	137	50
Patent No. 5514582	124	17.7	133	51
Sequence 57, Appl	265	17.7	133	52
Sequence 57, Appl	265	17.7	133	53
Sequence 2, Appl	764	17.7	133	54
Sequence 2, Appl	764	17.7	133	55
Sequence 46, Appl	177	17.2	129	56
Sequence 2, Appl	248	16.9	127	57
Sequence 4, Appl	266	16.9	127	58
Sequence 1, Appl	326	16.8	126	59
Patent No. 5514582	128	16.8	126	60
Patent No. 5514582	74	16.7	125	61
Sequence 28, Appl	76	16.7	125	62
Sequence 28, Appl	181	16.6	124	63
Patent No. 5478464	574	16.4	123	64
Sequence 4, Appl	610	16.0	120	65
Sequence 19, Appl	610	16.0	120	66
Sequence 89, Appl	610	16.0	120	67
Patent No. 5217870	610	16.0	120	68
Sequence 9, Appl	484	15.7	118	69
Sequence 9, Appl	484	15.7	118	70
Patent No. 5514582	126	15.4	115	71
Sequence 4, Appl	830	15.2	114	72
Sequence 2, Appl	830	15.2	114	73
Sequence 2, Appl	830	15.0	112	74
Sequence 24, Appl	216	14.8	111	75
Sequence 37, Appl	216	14.7	110	76
Sequence 48, Appl	123	14.6	110	77
Patent No. 5514582	123	14.5	109	78
Patent No. 5514582	127	14.5	109	79
Sequence 12, Appl	62	14.1	106	80
Sequence 12, Appl	62	14.1	106	81
Sequence 12, Appl	62	14.1	106	82
Patent No. 5514582	128	14.0	104	83
Patent No. 5514582	128	14.0	104	84
Patent No. 5514582	128	14.0	104	85
Sequence 13, Appl	62	13.8	103	86
Sequence 13, Appl	62	13.8	103	87
Sequence 13, Appl	62	13.8	103	88
Sequence 13, Appl	62	13.8	103	89
Sequence 13, Appl	62	13.8	103	90
Sequence 13, Appl	62	13.8	103	91
Sequence 13, Appl	62	13.8	103	92
Sequence 13, Appl	62	13.8	103	93
Sequence 13, Appl	62	13.8	103	94
Sequence 13, Appl	62	13.8	103	95
Sequence 13, Appl	62	13.8	103	96
Sequence 13, Appl	62	13.8	103	97
Sequence 13, Appl	62	13.8	103	98
Sequence 13, Appl	62	13.8	103	99
Sequence 13, Appl	62	13.8	103	100

ALIGNMENTS

RESULT 1
 Patent No. 5,256,642
 APPLICANT: FEAR, N., DOUGLAS T.; KLICKSTEIN, LLOYD B.; WONG,
 WINNIE W.; CARSON, GERALD R.; CONNIN, MICHAEL F.; LIP, STEPHEN
 H.; MAKRIDES, SAVVAS; MARSH, HENRY C., JR.
 TITLE OF INVENTION: COMPOSITIONS OF SOLUBLE COMPLEMENT
 RECEPTOR 1 (CR1) AND A THROMBOLYTIC AGENT, AND THE METHODS OF
 USE THEREOF
 NUMBER OF SEQUENCES: 40
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/588,128
 FILING DATE: 24 SEP 1999
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: 412,745
 FILING DATE: 26 SEP 1989
 APPLICATION NUMBER: 412,865
 FILING DATE: 03 APR 1989
 APPLICATION NUMBER: 176,532
 FILING DATE: 01 APR 1988
 SEQ ID NO: 6
 LENGTH: 1466
 5,256,642 6
 Query Match: 29.28; Score 219; DB 6; Length 1466;
 Best Local Similarity: 46.78; Prod. No. 2,40 19;
 Matches 51; Conservative 17; Mismatches 55; Indels 16; Gaps 6;

QY 2 ISGSGHPIHNGKSYYS --LPIAVGVIVYSQ SGRFLDKSLDITPKDY 52
 DB 1015 ISGSGHPIHNGKSYYS --LPIAVGVIVYSQ SGRFLDKSLDITPKDY 1072
 QY 53 DGTWKDPAPKCEYENKYSCEPIVIGVGRKIPGSTP-YRHLSTVEN-KTNSMKNKSV 111
 DB 1073 VVWWSPPRCSTNKR -CTAPEVENALRVGNSTSTETETTRGQPPVWWSHIV 1129
 QY 112 WQANNMWGPTRIPCVSV 140
 DB 1130 QQINGRWGP-KLPICSRV 1147
 RESULT 2
 Patent No. 5,472,939
 APPLICANT: FEAR, N., DOUGLAS T.; KLICKSTEIN, LLOYD B.; WONG,
 WINNIE W.; CARSON, GERALD R.; CONNIN, MICHAEL F.; LIP, STEPHEN
 H.; MAKRIDES, SAVVAS; MARSH, HENRY C., JR.
 TITLE OF INVENTION: METHOD OF TREATING COMPLEMENT
 MEDIATED DISORDERS
 NUMBER OF SEQUENCES: 40
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/138,825
 FILING DATE: 19 OCT 1994
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: 588,128
 FILING DATE: 24 SEP 1999
 APPLICATION NUMBER: 412,745
 FILING DATE: 01 APR 1988
 SEQ ID NO: 6
 LENGTH: 1466
 5,472,939 7
 Query Match: 29.28; Score 219; DB 6; Length 1466;
 Best Local Similarity: 46.78; Prod. No. 2,40 19;

Matches 51; Conservative 17; Mismatches 55; Indels 16; Gaps 6;
 QY 2 ISGSGHPIHNGKSYYS --LPIAVGVIVYSQ SGRFLDKSLDITPKDY 52
 DB 1015 ISGSGHPIHNGKSYYS --LPIAVGVIVYSQ SGRFLDKSLDITPKDY 1072
 QY 53 DGTWKDPAPKCEYENKYSCEPIVIGVGRKIPGSTP-YRHLSTVEN-KTNSMKNKSV 111
 DB 1073 VVWWSPPRCSTNKR -CTAPEVENALRVGNSTSTETETTRGQPPVWWSHIV 1129
 QY 112 WQANNMWGPTRIPCVSV 140
 DB 1130 QQINGRWGP-KLPICSRV 1147
 RESULT 3
 Patent No. 5,256,642
 APPLICANT: FEAR, N., DOUGLAS T.; KLICKSTEIN, LLOYD B.; WONG,
 WINNIE W.; CARSON, GERALD R.; CONNIN, MICHAEL F.; LIP, STEPHEN
 H.; MAKRIDES, SAVVAS; MARSH, HENRY C., JR.
 TITLE OF INVENTION: COMPOSITIONS OF SOLUBLE COMPLEMENT
 RECEPTOR 1 (CR1) AND A THROMBOLYTIC AGENT, AND THE METHODS OF
 USE THEREOF
 NUMBER OF SEQUENCES: 40
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/588,128
 FILING DATE: 24 SEP 1999
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: 412,745
 FILING DATE: 26 SEP 1989
 APPLICATION NUMBER: 412,865
 FILING DATE: 03 APR 1989
 APPLICATION NUMBER: 176,532
 FILING DATE: 01 APR 1988
 SEQ ID NO: 5
 LENGTH: 1537
 5,256,642 5
 Query Match: 29.28; Score 219; DB 6; Length 1537;
 Best Local Similarity: 46.78; Prod. No. 2,40 19;
 Matches 51; Conservative 17; Mismatches 55; Indels 16; Gaps 6;
 QY 2 ISGSGHPIHNGKSYYS --LPIAVGVIVYSQ SGRFLDKSLDITPKDY 52
 DB 1015 ISGSGHPIHNGKSYYS --LPIAVGVIVYSQ SGRFLDKSLDITPKDY 1072
 QY 53 DGTWKDPAPKCEYENKYSCEPIVIGVGRKIPGSTP-YRHLSTVEN-KTNSMKNKSV 111
 DB 1073 VVWWSPPRCSTNKR -CTAPEVENALRVGNSTSTETETTRGQPPVWWSHIV 1129
 QY 112 WQANNMWGPTRIPCVSV 140
 DB 1130 QQINGRWGP-KLPICSRV 1147
 RESULT 4
 Patent No. 5,472,939
 APPLICANT: FEAR, N., DOUGLAS T.; KLICKSTEIN, LLOYD B.; WONG,
 WINNIE W.; CARSON, GERALD R.; CONNIN, MICHAEL F.; LIP, STEPHEN
 H.; MAKRIDES, SAVVAS; MARSH, HENRY C., JR.
 TITLE OF INVENTION: METHOD OF TREATING COMPLEMENT
 MEDIATED DISORDERS
 NUMBER OF SEQUENCES: 40
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/138,825
 FILING DATE: 19 OCT 1994
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: 588,128
 FILING DATE: 24 SEP 1999
 APPLICATION NUMBER: 412,745
 FILING DATE: 26 SEP 1989

DB 1662 QUONERWOD KIPHSRV 1649

REGFILE 8

6472669 2

PATENT NO. 512969

APPLICANT: FERRON, DOUGLAS L.; KILICKSTEIN, LEONARD B.; WONG, JENNELL W.; CARSON, GERALD K.; CONNING, MICHAEL F.; LIP, STEPHEN H.; MAKRESIS, SAVVAS; MARSH, HENRY C., JR.
TITLE OF INVENTION: METHOD OF TREATING COMPLEMENT RELATED DISORDERS

NUMBER OF SEQUENCES: 40

CURRENT APPLICATION DATA:

APPLICATION NUMBER: 03/067108 825

FILING DATE: 19 OCT 1994

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 5808128

FILING DATE: 24 SEP 1990

APPLICATION NUMBER: 112545

FILING DATE: 26 SEP 1989

APPLICATION NUMBER: 352865

FILING DATE: 03 APR 1989

APPLICATION NUMBER: 176542

FILING DATE: 01 APR 1989

DB 1662 QUONERWOD KIPHSRV 1649

LENGTH: 2039

6472669 2

Query Match 29.2% Score 219 DB 63 Length 2039

Best Local Similarity 30.2% Pred. No. 340 152

Mismatches 51 Conservative 17 Mismatches 55 Indels 16 Gaps 6

QY 4 TGGGCTTTRKRLVY TLAVGVVYVS* SCHKLGGZSLTTRKRV 52

DB 1517 ISCTPTTSG DYSSNRTSGVWVYQ HIGDQQLTLLGVSTYISKDQ 1594

QY 53 DEDWDKAPKTYENKSSSTPTVYVYKESSEYHRSVGVATKINSSNKSRY 111

DB 1596 VVWSSTPTPTSTNK CTAEVENATVGNKSESTLTKRQQLVWGSHTV 1641

QY 112 WQANNWGDREPTVSV 140

DB 1662 QUONERWOD KIPHSRV 1649

REGFILE 9

US 09-834-309-4

Sequence 2, Application in US/Europe

PATENT NO. 5187268

GENERAL INFORMATION

APPLICANT: Kowal, Gil-ah

APPLICANT: Moss, Bernard

TITLE OF INVENTION: Synthetic, Anti Complement Protein and

TITLE OF INVENTION: The Gene Encoding Same

NUMBER OF SEQUENCES: 4

CORRESPONDENCE ADDRESS:

ADDRESSEE: Townsend and Townsend

STREET: One Market Plaza, Steuart Tower, Suite 2000

CITY: San Francisco

STATE: California

COUNTRY: USA

ZIP: 94105

COMPUTER READABLE FORM

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC DOS/MS DOS

SOFTWARE: Patent in Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: 03/067108 825

FILING DATE: 1994/01

CLASSIFICATION: 580

AGENCY/AGENT INFORMATION:

NAME: Weber, Kenneth A.

REGISTRATION NUMBER: 41,677

REFERENCE/SEQ ID NO: 15280 9

TELECOMMUNICATION INFORMATION:

TELEPHONE: 415 543 9600

TELEFAX: 415 543 5043

INFORMATION FOR SEQ ID NO: 2:

SEQUENCE CHARACTERISTICS:

LENGTH: 263 amino acids

TYPE: AMINO ACID

TOPOLOGY: linear

MOLECULE TYPE: protein

US-07-906 984 2

Query Match 29.5% Score 214 DB 13 Length 263

Best Local Similarity 35.9% Pred. No. 840 16

Mismatches 46 Conservative 15 Mismatches 55 Indels 12 Gaps 5

QY 2 DSGSDPTTNRKSYSTPVAIVYVSSESTLTKRQQLVWGSHTV 61

DB 146 VKQPTPTTRGRGRYDPTDGSVIVYVSSESTLTKRQQLVWGSHTV 190

QY 62 KQYVTPYVSSTLPAVLSYFETLSLYSHAVLAVKIN 85 82 VVWVLSHHW 120

DB 200 TQVAV--KQHPDPTSNVYSSGPKKSYSVNWDKRVKVVAVSSSSSTPTSNW 255

QY 121 PTKLTCV 128

DB 256 PFLPKCV 262

RESULT 10

US 08-356-461-41

Sequence 41, Application 95/08456461

PATENT NO. 5843889

GENERAL INFORMATION:

APPLICANT: Smith, Richard A.G.

APPLICANT: Bold, Ian

APPLICANT: Freeman Mary A.

APPLICANT: Mossakowska, Barbara E.L.

TITLE OF INVENTION: No. 58438894 Compounds

NUMBER OF SEQUENCES: 41

CORRESPONDENCE ADDRESS:

ADDRESSEE: SmithKline Beecham Corporate Intellectual Property

STREET: P.O. Box 1549

CITY: King of Prussia

STATE: Pennsylvania

COUNTRY: USA

ZIP: 19406

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC DOS/MS DOS

SOFTWARE: Patent in Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: 03/087356, 461

FILING DATE: 04 JUL 1995

CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

NAME: Jarvis, Herbert H.

REGISTRATION NUMBER: 41,171

REFERENCE/SEQ ID NO: 130424

TELECOMMUNICATION INFORMATION:

TELEPHONE: (610) 270 5019

TELEFAX: (610) 270 5090

INFORMATION FOR SEQ ID NO: 41:

SEQUENCE CHARACTERISTICS:

LENGTH: 34 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: peptide

FRAGMENT TYPE: Internal

US 08-356-461-41

Query Match
Best Local Similarity 26.4% Score 198; DB 2; Length 133;
Matches 51; Conservative 14; Mismatches 57; Indels 12; Gaps 6;

QY 2 ISCGSPHLLNCR-LSYSTPIAGCTVIRSC---SG---TFRLGKSKSLCTTKRKVK 54
DB 3 ITCGLPTITINDGFTSTNRENFHSGSVVTVKCNCGSKRKVFELVGEPSIYCTSNIDQVG 62

QY 55 TWKDPAPKCEYFNKYSSCPETIVEGKYKIGAS-TPYKHGDSVTEAKTNFSMNKNKSVWC 113
DB 63 IWSGPAPOCTIPNK---CTPPNVENGIIIVSNRSLFSLNFVVEPPQPGFVVMKGPFRVK 119

QY 114 QANNMWGPTRLPTC 127
DB 120 QALNKWEP-ELFSC 132

RESULT 11
US-08-769-967A-29
; Sequence 31, Application US/08769967A
; Patent No. 5859224
; GENERAL INFORMATION:
; APPLICANT: Mossakowska, Danuta E.I.
; APPLICANT: Dadd, Ian
; APPLICANT: Freeman, Anne Mary
; TITLE OF INVENTION: Soluble CR1 Derivatives
; NUMBER OF SEQUENCES: 33
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SmithKline Beecham Corporate Intellectual Property
; STREET: P.O. Box 1539
; CITY: King of Prussia
; STATE: Pennsylvania
; COUNTRY: USA
; ZIP: 19406
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/769,967A
; FILING DATE: 15 May 1995
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/440,569
; FILING DATE: 15 May 1995
; ATTORNEY/AGENT INFORMATION:
; NAME: King, William T.
; REGISTRATION NUMBER: 40,954
; REFERENCE/DOCKET NUMBER: P-042002
; TELEPHONE: (610) 270-5464
; TELEFAX: (610) 270-5090
; INFORMATION FOR SEQ ID NO: 31:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 133 amino acids
; TYPE: amino acid
; TOPOLOGY: Linear
; MOLECULE TYPE: peptide
; FRAGMENT TYPE: internal
US-08-769-967A-31

Query Match
Best Local Similarity 26.4% Score 198; DB 2; Length 133;
Matches 51; Conservative 14; Mismatches 57; Indels 12; Gaps 6;

QY 2 ISCGSPHLLNCR-LSYSTPIAGCTVIRSC---SG---TFRLGKSKSLCTTKRKVK 54
DB 3 ITCGLPTITINDGFTSTNRENFHSGSVVTVKCNCGSKRKVFELVGEPSIYCTSNIDQVG 62

QY 55 TWKDPAPKCEYFNKYSSCPETIVEGKYKIGAS-TPYKHGDSVTEAKTNFSMNKNKSVWC 113
DB 63 IWSGPAPOCTIPNK---CTPPNVENGIIIVSNRSLFSLNFVVEPPQPGFVVMKGPFRVK 119

DB 63 IWSGPAPOCTIPNK---CTPPNVENGIIIVSNRSLFSLNFVVEPPQPGFVVMKGPFRVK 119

QY 114 QANNMWGPTRLPTC 127
DB 120 QALNKWEP-ELFSC 132

RESULT 12
US-08-356-361-29
; Sequence 29, Application US/08356361
; Patent No. 5833989
; GENERAL INFORMATION:
; APPLICANT: Smith, Richard A.G.
; APPLICANT: Dadd, Ian
; APPLICANT: Freeman Mary A.
; APPLICANT: Mossakowska, Danuta E.I.
; TITLE OF INVENTION: No. 5833989el (comounds
; NUMBER OF SEQUENCES: 31
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SmithKline Beecham Corporate Intellectual Property
; STREET: P.O. Box 1539
; CITY: King of Prussia
; STATE: Pennsylvania
; COUNTRY: USA
; ZIP: 19406
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/356,361
; FILING DATE: 03-Jul-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Jctvls, Herbert H.
; REGISTRATION NUMBER: 31,171
; REFERENCE/DOCKET NUMBER: P30423
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (610) 270-5019
; TELEFAX: (610) 270-5090
; INFORMATION FOR SEQ ID NO: 29:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 254 amino acids
; TYPE: amino acid
; TOPOLOGY: Linear
; MOLECULE TYPE: peptide
; FRAGMENT TYPE: N-terminal
US-08-356-361-29

Query Match
Best Local Similarity 26.4% Score 198; DB 2; Length 254;
Matches 51; Conservative 14; Mismatches 57; Indels 12; Gaps 6;

QY 2 ISCGSPHLLNCR-LSYSTPIAGCTVIRSC---SG---TFRLGKSKSLCTTKRKVK 54
DB 3 ITCGLPTITINDGFTSTNRENFHSGSVVTVKCNCGSKRKVFELVGEPSIYCTSNIDQVG 62

QY 55 TWKDPAPKCEYFNKYSSCPETIVEGKYKIGAS-TPYKHGDSVTEAKTNFSMNKNKSVWC 113
DB 184 IWSGPAPOCTIPNK---CTPPNVENGIIIVSNRSLFSLNFVVEPPQPGFVVMKGPFRVK 240

QY 114 QANNMWGPTRLPTC 127
DB 241 QALNKWEP-ELFSC 253

RESULT 13
US-08-769-967A-29
; Sequence 29, Application US/08769967A
; Patent No. 5859224
; GENERAL INFORMATION:
; APPLICANT: Mossakowska, Danuta E.I.

? FILING DATE: 15-May-1995
? ATTORNEY/AGENT INFORMATION:
? NAME: Kind, William T.
? REGISTRATION NUMBER: 30,954
? REFERENCE/DOCKET NUMBER: P4042402
? TELECOMMUNICATION INFORMATION:
? TELEPHONE: (610) 270-5464
? TELEFAX: (610) 270-5090
? INFORMATION FOR SEQ ID NO: 30:
? SEQUENCE CHARACTERISTICS:
? LENGTH: 254 amino acids
? TYPE: amino acid
? TOPOLOGY: linear
? MOLECULE TYPE: peptide
? FRAGMENT TYPE: N-terminal
US 09 967A 40

Query Match: 25.04; Score 195; DB 2; Length 254;
Best Local Similarity 98.16; Pred. No. 9,66-14;
Matches 51; Conservative 13; Mismatches 58; Indels 12; Gaps 6;

QY 2 ISGSGPPILNGR-LSYYSPIAVGTIVYSC- -SG- - - - -TERLIGKSLICITIKKVDG 54
DB 124 IEGGLPTIINGUFTSTNPFNHYGVVYTPGNPISGPKVFELVGEPIYCTSDNDQVG 183
QY 55 TWKPAKPEYENKYSCEPIIPVGGYKIRG-STYPRKHSVTFK-TNFSMNGKNSVMC 113
DB 184 IWSGLAPQVLIUNK- -CTFNVENGLIVASINRSLSINEVEPRCOIGFVMKGPDRVKC 240
QY 114 QANNMGPRRLPTCV 127
DB 241 QALNKWEP ELPSC 253

RESULT 16
US 09 961 234B 2
? Sequence 2, Application US/09081234B
? Patent No. 6,270,997
? GENERAL INFORMATION:
? APPLICANT: TOYOMURA, KOJI
? APPLICANT: MURAKAMI, HIROSHI
? APPLICANT: SHIGEMURA, TAMOTSU
? TITLE OF INVENTION: DNA ENCODING A PORCINE COMPLEMENT
? NUMBER OF SEQUENCES: 2
? CORRESPONDENCE ADDRESS:
? ADDRESSEE: BIRCH, STEWART, KOLASCH & BIRCH
? STREET: PO BOX 747
? CITY: FALLS CHURCH
? STATE: VA
? COUNTRY: USA
? ZIP: 22040-0747
? COMPUTER READABLE FORM:
? MEDIUM TYPE: Floppy disk
? COMPUTER: IBM PC compatible
? OPERATING SYSTEM: PC DOS/MS DOS
? SOFTWARE: Patent In Release #1.0, Version #1.40
? CURRENT APPLICATION DATA:
? APPLICATION NUMBER: US/09/961,234B
? FILING DATE: 12-DEC-1997
? CLASSIFICATION:

? ATTORNEY/AGENT INFORMATION:
? NAME: MURPHY JR, GERALD M.
? REGISTRATION NUMBER: 28,977
? REFERENCE/DOCKET NUMBER: 2520-111P
? TELECOMMUNICATION INFORMATION:
? TELEPHONE: 703 205 8000
? TELEFAX: 703 205 8050
? INFORMATION FOR SEQ ID NO: 2:
? SEQUENCE CHARACTERISTICS:
? LENGTH: 463 amino acids
? TYPE: amino acid
? TOPOLOGY: linear

? MOLECULE TYPE: protein
US-09-981-234B-2

Query Match: 25.38; Score 193; DB 4; Length 463;
Best Local Similarity 33.68; Pred. No. 3,40-14;
Matches 45; Conservative 20; Mismatches 53; Indels 16; Gaps 7;

QY 2 ISGSGPPILNGR-LSYYSPIAVGTIVYSC- -ST- - - - -ERLIGKSLICITIKKVDG 53
DB 169 IICKPPPEIPNCKYTNKSHKDVFNENVVYVYSGTGHDFPSWGESSEFCCKDF- - 224
QY 56 WDKPAKPEYENKYSCEPIIPVGGYKIRG-STYPRKHSVTFK-TNFSMNGKNSVMC 114
DB 225 WSDPPEC- -KVKCPYPVPVNGEIVSGHSFYKAEVVPKCNAGFTILRGRTIVG 280
QY 115 ANNMGGPRRLPTCV 128
DB 281 ANSTWEP EMPQCI 293

RESULT 17
US-09-139-195-2

? Sequence 2, Application US/08139195
? Patent No. 6,218,520
? GENERAL INFORMATION:
? APPLICANT: Atkinson, John P.
? TITLE OF INVENTION: RECOMBINANTLY PRODUCED HUMAN MEMBRANE
? TITLE OF INVENTION: COFACTOR PROTEIN (MCP)
? NUMBER OF SEQUENCES: 2

? CORRESPONDENCE ADDRESS:
? ADDRESSEE: Patrea L. Pabst
? STREET: 1100 Peachtree Street, SW
? CITY: Atlanta
? STATE: Georgia
? COUNTRY: USA
? ZIP: 30309-4530

? COMPUTER READABLE FORM:
? MEDIUM TYPE: Floppy disk
? COMPUTER: IBM PC compatible
? OPERATING SYSTEM: PC DOS/MS DOS
? SOFTWARE: Patent In Release #1.0, Version #1.25
? CURRENT APPLICATION DATA:

? APPLICATION NUMBER: US/08/139,195
? FILING DATE: 20-OCT-1993
? CLASSIFICATION: 435

? PRIOR APPLICATION DATA:
? APPLICATION NUMBER: US 07/948,350

? FILING DATE: 21-SEP-1992
? PRIOR APPLICATION DATA:

? APPLICATION NUMBER: US 07/484,211
? FILING DATE: 21-JUL-1989

? ATTORNEY/AGENT INFORMATION:
? NAME: Pabst, Patrea L.

? REGISTRATION NUMBER: 31,284
? REFERENCE/DOCKET NUMBER: W010040N

? TELECOMMUNICATION INFORMATION:
? TELEPHONE: (404)-815-6508

? TELEFAX: (404)-815-6555
? INFORMATION FOR SEQ ID NO: 2:

? SEQUENCE CHARACTERISTICS:
? LENGTH: 484 amino acids
? TYPE: amino acid

? STRANDEDNESS: single
? TOPOLOGY: linear

? MOLECULE TYPE: protein
? HYPOTHETICAL: NO

? ANTI-SENSE: NO
? FRAGMENT TYPE: N-terminal

? ORIGINAL SOURCE:
? ORGANISM: Human Membrane CoFactor Protein (MCP)

US-09-139-195 2

Query Match

21.38; Score 160; DB 4; Length 484;

CORRESPONDENCE ADDRESS:
 ADDRESSEE: Fish & Richardson P.C.
 STREET: 225 Franklin Street
 CITY: Boston
 STATE: Massachusetts
 COUNTRY: U.S.A.
 ZIP: 02110 2804
 COMPUTER READABLE FORM:
 MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
 COMPUTER: IBM PS/2 Model 50Z or 55SX
 OPERATING SYSTEM: MS DOS (Version 5.0)
 SOFTWARE: WordPerfect (Version 5.1)
 CURRENT APPLICATION DATA:
 FILING DATE: 22 SEP 1994
 APPLICATION NUMBER: US/09/410,416A
 FILING DATE: 22 SEP 1994
 CLASSIFICATION: 435
 ATTORNEY/AGENT INFORMATION:
 NAME: Paul T. Clark
 REGISTRATION NUMBER: 40,162
 REFERENCE/DOCKET NUMBER: 06180/005001
 TELEPHONE: (617) 542-5070
 TELEFAX: (617) 542-8906
 TELEX: 200154
 INFORMATION FOR SEQ ID NO: 13:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 254 amino acids
 TYPE: amino acid
 TOPOLOGY: linear
 SURNAME: not relevant
 MOLECULE TYPE: protein
 US-09-410A-13

Query Match 21.2%, Score 159; DB 1; length 254;
 Best Local Similarity 27.9%; Pred. No. 8.6e-10;
 Matches 48; Conservative 25; Mismatches 57; Indels 16; Gaps 6;

QY 2 ISGSPHPLNGRISYSYSTPI AVCTVIRYSCS-----GTERLIGERSKSLCTTKDKVWGI 55
 DB 126 VICTPIPKIKCKHTESEVVEFYLDVAVTSCDAPGHPDPSLIGESTIYV-----GINSV 181
 QY 56 WDKFAKCEYENKYSCEPEIVFGYKIRG STYRHCQAVTFACKTNFNMNKNKSWVQ 114
 DB 182 WSKAAEC---KVKCRFPVWVNGKQISGRKFPYKATVMEFCDKGYLDGSDITVCD 237
 QY 115 ANNMWPTPLTCVSV 130
 DB 238 SNSTWPP VPKCLV 252

RESULT 21
 US-09-888-171-13
 Sequence 13, Application US/0888171
 Patent No. 5851528
 GENERAL INFORMATION:
 APPLICANT: Zome Corp., K
 APPLICANT: Higgins, Paul J.
 APPLICANT: Yeh, C. Grace
 TITLE OF INVENTION: METHODS OF INHIBITING COMPLEMENT
 TITLE OF INVENTION: ACTIVATION
 NUMBER OF SEQUENCES: 19
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Fish & Richardson, P.C.
 STREET: 225 Franklin Street
 CITY: Boston
 STATE: MA
 COUNTRY: US
 ZIP: 02110 2804
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Diskette
 COMPUTER: IBM Compatible
 OPERATING SYSTEM: Windows
 SOFTWARE: FastSeq for Windows Version 2.0

CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/888,171
 FILING DATE: 03-JUL-1997
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: 08/310,416
 FILING DATE: 22-SEP-1994
 ATTORNEY/AGENT INFORMATION:
 NAME: Freeman, John W.
 REGISTRATION NUMBER: 29,066
 REFERENCE/DOCKET NUMBER: 06180/0,5012
 TELEPHONE: 617/542-507
 TELEFAX: 617/542-890
 TELEX: 200154
 INFORMATION FOR SEQ ID NO: 13:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 254 amino acids
 TYPE: amino acid
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 US-08-888-171-13

Query Match 21.2%, Score 159; DB 2; length 254;
 Best Local Similarity 27.9%; Pred. No. 8.6e-10;
 Matches 38; Conservative 25; Mismatches 57; Indels 16; Gaps 6;

QY 2 ISGSPHPLNGRISYSYSTPI AVCTVIRYSCS-----GTERLIGERSKSLCTTKDKVWGI 55
 DB 126 VICTPIPKIKCKHTESEVVEFYLDVAVTSCDAPGHPDPSLIGESTIYV-----GINSV 181
 QY 56 WDKFAKCEYENKYSCEPEIVFGYKIRG STYRHCQAVTFACKTNFNMNKNKSWVQ 114
 DB 182 WSKAAEC---KVKCRFPVWVNGKQISGRKFPYKATVMEFCDKGYLDGSDITVCD 247
 QY 115 ANNMWPTPLTCVSV 130
 DB 238 SNSTWPP VPKCLV 252

RESULT 22
 US-08-435-149-1
 Sequence 1, Application US/08435149
 Patent No. 5836402
 GENERAL INFORMATION:
 APPLICANT: INNIS, MICHAEL A.
 APPLICANT: ZAKOR, ISABEL
 APPLICANT: GREASEY, ABRAHAM
 TITLE OF INVENTION: CHIMERIC MCP AND DAF PROTEINS WITH CELL
 SURFACE LOCALIZING DOMAIN
 NUMBER OF SEQUENCES: 26
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: CHIRON CORPORATION
 STREET: INTELLECTUAL PROPERTY - 8440, P.O. BOX 8097
 CITY: EMERYVILLE
 STATE: CALIFORNIA
 COUNTRY: U.S.A.
 ZIP: 94662-8097
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/435,149
 FILING DATE: 05-MAY-1995
 CLASSIFICATION: 530
 ATTORNEY/AGENT INFORMATION:
 NAME: SAVERIDE, PAUL B.
 REGISTRATION NUMBER: 36,914
 REFERENCE/DOCKET NUMBER: 0989,001
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (510) 602-2585
 TELEFAX: (510) 655-3542

TELEX: N/A
 INFORMATION FOR SEQ ID NO: 1:
 LENGTH: 294 amino acids
 TYPE: amino acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 US 09 834 309 4

Query Match 21.2% Score 159; DB 1; Length 294;

Best local similarity 27.9% Pred. No. 16 09;

Matches 48 Conservative 25; Mismatches 67; Indels 16; Gaps 67

QY 2 ISGSGPPTLNGKSYSTPLAVGVIVSYSS -GIFRLGKSLGCLIKKVKVMT 55
 DB 126 VLTLPPTKFNKHIFSEVEVEFYLLAVYSYDPAIDGDPSEIIGSTLYG GNSV 181
 QY 56 WKELAKFYEYKSYSTPLAVGVIVSYSS STPTKIGDSVIFA'KINFSMGNKSVWQ 114
 DB 182 WSRAPF-----KVKRQFVVVNGKQISGCKKFKYKATVMEKTKGYLNGSDIIVGD 237
 QY 115 ANNMWGPTRPFCVSV 140
 DB 238 SNSTWDPVPKCLKV 252

RESULT 24

US 09 834 309 4

Sequence 16, Application US/09/041416A

Patent No. 567546

GENERAL INFORMATION:

APPLICANT: Jono-Lou, K et al.

TITLE OF INVENTION: CHEMICAL PROTEINS WHICH BLOCK

TITLE OF INVENTION: COMPLEMENT ACTIVATION

NUMBER OF SEQUENCES: 19

CORRESPONDENCE ADDRESS:

ADDRESSEE: Fish & Richardson P.C.

STREET: 225 Franklin Street

CITY: Boston

STATE: Massachusetts

COUNTRY: U.S.A.

ZIP: 02110-2804

COMPUTER READABLE FORM:

MEDIUM TYPE: 3.5" Diskette, 1.44 MB

COMPUTER: IBM PS/2 Model 55SX

OPERATING SYSTEM: MS DOS (Version 5.0)

SOFTWARE: WordPerfect (Version 5.1)

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/041416A

FILING DATE: 22 SEP 1994

CLASSIFICATION: 4.05

ATTORNEY/AGENT INFORMATION:

NAME: Paul T. Clark

REGISTRATION NUMBER: 40162

REFERENCE/REGISTER NUMBER: 06180208001

TELEPHONE: (617) 542-5000

TELEFAX: (617) 542-8900

TELEX: 200154

INFORMATION FOR SEQ ID NO: 1:

SEQUENCE CHARACTERISTICS:

LENGTH: 294 amino acids

TYPE: amino acid

STRANDEDNESS: not determined

TOPOLOGY: linear

MOLECULE TYPE: protein

US 09 834 309 4

Query Match 21.2% Score 159; DB 1; Length 294;

Best local similarity 27.9% Pred. No. 16 09;

Matches 48 Conservative 25; Mismatches 67; Indels 16; Gaps 67

QY 2 ISGSGPPTLNGKSYSTPLAVGVIVSYSS -GIFRLGKSLGCLIKKVKVMT 55
 DB 126 VLTLPPTKFNKHIFSEVEVEFYLLAVYSYDPAIDGDPSEIIGSTLYG GNSV 181
 QY 56 WKELAKFYEYKSYSTPLAVGVIVSYSS STPTKIGDSVIFA'KINFSMGNKSVWQ 114
 DB 182 WSRAPF-----KVKRQFVVVNGKQISGCKKFKYKATVMEKTKGYLNGSDIIVGD 237
 QY 115 ANNMWGPTRPFCVSV 140
 DB 238 SNSTWDPVPKCLKV 252

RESULT 24

US 09 888-171 16

Sequence 16, Application US/09/088171

Patent No. 5651528

GENERAL INFORMATION:

APPLICANT: Jono-Lou, K et al.

TITLE OF INVENTION: CHEMICAL PROTEINS WHICH BLOCK

TITLE OF INVENTION: COMPLEMENT ACTIVATION

NUMBER OF SEQUENCES: 19

CORRESPONDENCE ADDRESS:

ADDRESSEE: Fish & Richardson P.C.

STREET: 225 Franklin Street

CITY: Boston

STATE: MA

COUNTRY: US

ZIP: 02110-2804

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette

COMPUTER: IBM Compatible

OPERATING SYSTEM: Windows95

SOFTWARE: FastSeq for Windows Version 2.0

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/088171

FILING DATE: 04-JUL-1997

PREVIOUS APPLICATION DATA:

APPLICATION NUMBER: 08/410,416

FILING DATE: 22 SEP-1994

ATTORNEY/AGENT INFORMATION:

NAME: Freeman, John W.

REGISTRATION NUMBER: 29,066

REFERENCE/REGISTER NUMBER: 00100,000002

TELEPHONE: 617/542-5557

TELEFAX: 617/542-8900

TELEX: 200154

INFORMATION FOR SEQ ID NO: 16:

SEQUENCE CHARACTERISTICS:

LENGTH: 294 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

US 09 888-171-16

Query Match 21.2% Score 159; DB 2; Length 294;

Best local similarity 27.9% Pred. No. 16 09;

Matches 48 Conservative 25; Mismatches 67; Indels 16; Gaps 67

QY 2 ISGSGPPTLNGKSYSTPLAVGVIVSYSS -GIFRLGKSLGCLIKKVKVMT 55
 DB 126 VLTLPPTKFNKHIFSEVEVEFYLLAVYSYDPAIDGDPSEIIGSTLYG GNSV 181
 QY 56 WKELAKFYEYKSYSTPLAVGVIVSYSS STPTKIGDSVIFA'KINFSMGNKSVWQ 114
 DB 182 WSRAPF-----KVKRQFVVVNGKQISGCKKFKYKATVMEKTKGYLNGSDIIVGD 237
 QY 115 ANNMWGPTRPFCVSV 140
 DB 238 SNSTWDPVPKCLKV 252


```

QY 117 N MW GPTLEPTC 127
      1 1 1 1 1 1
DB 491 NDEGEWSGPP--DEC 503

RESULT 40
US 09 475 460A 32
: Sequence 42, Application US/09475460A
: Patent No. 6416253
: GENERAL INFORMATION:
: APPLICANT: Scott, Elizabeth
: TITLE OF INVENTION: EXPRESSION VECTORS, TRANSECTION SYSTEMS, AND METHOD OF
: FILE REFERENCE: 1527-004
: CURRENT FILING DATE: 1999-12-30
: NUMBER OF SEQ ID NOS: 32
: SOFTWARE: Patent In Ver. 2.0
: SEQ ID NO 42
: LENGTH: 611
: TYPE: PPT
: ORGANISM: CAR2
US 09 475 460A 32

Query Match
Best Local Similarity 21.2%; Score 159; DB 4; Length 611;
Matches 45; Conservative 17; Mismatches 50; Indels 23; Caps 9;

QY 4 SGGSPPLINGRISYSTP--IAVTVIRYSCGTFRLIGEKSLLCITKQVDGT---WD 57
      1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
DB 416 SCNPNGEIRNGQI---DWPGGILFCATISTSCNTGYKLFQGSTSSFCI---ISGSSVQWS 468

QY 58 KPAKCEYFNKYSSCPEDIVGGYKIKGSDP-YRHCUSVTFACKTNFNMNKNKSVWCOAN 116
      1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
DB 469 DPLPEC---REIYCPAPPQIDNGIIQGER-HYGRQSVTVACNKGFTMIGHSIYCTVN 524

QY 117 N MW GPTLEPTC 127
      1 1 1 1 1 1
DB 525 NDEGEWSGPP--DEC 547

RESULT 41
US 08 445 149-2
: Sequence 2, Application US/08445149
: Patent No. 5866402
: GENERAL INFORMATION:
: APPLICANT: INNIS, MICHAEL A.
: APPLICANT: ZAROB, ISABEL
: TITLE OF INVENTION: CHIMERIC MCP AND DAF PROTEINS WITH CELL
: FILE REFERENCE: SURFACE LOCALIZING DOMAIN
: NUMBER OF SEQUENCES: 26
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: CHIRON CORPORATION
: STREET: INTELLECTUAL PROPERTY
: CITY: EMERYVILLE
: STATE: CALIFORNIA
: COUNTRY: U.S.A.
: ZIP: 94662 8097
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: Patent In Release #1.0, Version #1.25
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/445,149
: FILING DATE: 05 MAY 1995
: CLASSIFICATION: 530
: ATTORNEY/AGENT INFORMATION:
: NAME: SAVERIDE, PAUL B.
: REGISTRATION NUMBER: 46,914

QY 117 N MW GPTLEPTC 127
      1 1 1 1 1 1
DB 491 NDEGEWSGPP--DEC 503

REFERENCE/DOI/KEY NUMBER: 0989 001
TELECOMMUNICATION INFORMATION:
TELEPHONE: (510) 601-2585
TELEFAX: (510) 655-3542
TELEX: N/A
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 324 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-435-149-2

Query Match
Best Local Similarity 21.1%; Score 158; DB 2; Length 323;
Matches 45; Conservative 17; Mismatches 60; Indels 24; Caps 9;

QY 4 SGGSPPLINGRISYSTP--IAVTVIRYSCGTFRLIGEKSLLCITKQVDGT---WD 57
      1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
DB 128 SCNPNGEIRNGQI---DWPGGILFCATISTSCNTGYKLFQGSTSSFCI---ISGSSVQWS 180

QY 58 KPAKCEYFNKYSSCPEDIVGGYKIKGSDP-YRHCUSVTFACKTNFNMNKNKSVWCOAN 116
      1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
DB 181 DPLPEC---REIYCPAPPQIDNGIIQGER-HYGRQSVTVACNKGFTMIGHSIYCTVN 246

QY 117 N MW GPTLEPTC 127
      1 1 1 1 1 1
DB 237 NDEGEWSGPP--DEC 249

RESULT 42
US-08-310-416A-11
: Sequence 14, Application US/08310416A
: Patent No. 5673546
: GENERAL INFORMATION:
: APPLICANT: Jone-Long Ko et al.
: TITLE OF INVENTION: CHIMERIC PROTEINS WHICH BLOCK
: FILE REFERENCE: COMPLEMENT ACTIVATION
: NUMBER OF SEQUENCES: 19
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Fish & Richardson P.C.
: STREET: 225 Franklin Street
: CITY: Boston
: STATE: Massachusetts
: COUNTRY: U.S.A.
: ZIP: 02110 2404
: COMPUTER READABLE FORM:
: MEDIUM TYPE: 3.5" Diskette, 1.44 MB
: COMPUTER: IBM PS/2 Model 502 or 505
: OPERATING SYSTEM: MS-DOS (Version 5.0)
: SOFTWARE: WordPerfect (Version 5.1)
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/310,416A
: FILING DATE: 22-SEP-1994
: CLASSIFICATION: 435
: ATTORNEY/AGENT INFORMATION:
: NAME: Paul T. Clark
: REGISTRATION NUMBER: 30,162
: REFERENCE/DOI/KEY NUMBER: 06180/005001
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (617) 542-5070
: TELEFAX: (617) 542-8906
: TELEX: 200154
: INFORMATION FOR SEQ ID NO: 14:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 324 amino acids
: TYPE: amino acid
: STRANDEDNESS: not relevant
: TOPOLOGY: linear
: MOLECULE TYPE: protein
US-08-310-416A-14

```


RESULT 5
US-08-456-461-27
Sequence 27, Application US/08456461
Patent No. 5833484
GENERAL INFORMATION:
APPLICANT: Smith, Richard A.G.
APPLICANT: Dodd, Ian
APPLICANT: Freeman Mary A.
APPLICANT: Mossakowska, Barbara E.L.
TITLE OF INVENTION: No. 5833484ol Compounds
NUMBER OF SEQUENCES: 31
CORRESPONDENCE ADDRESS:
ADDRESSEE: SmithKline Beecham Corporate Intellectual Property
STREET: P.O. Box 1549
CITY: King of Prussia
STATE: Pennsylvania
COUNTRY: USA
ZIP: 19406
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/456,461
FILING DATE: 03 Jul 1995
CLASSIFICATION: 445
ATTORNEY/AGENT INFORMATION:
NAME: Jervis, Herbert H.
REGISTRATION NUMBER: 4,171
REFERENCE/DOCKET NUMBER: P40424
TELECOMMUNICATION INFORMATION:
TELEPHONE: (610) 270-5019
TELEFAX: (610) 270-5090
INFORMATION FOR SEQ ID NO: 27:
SEQUENCE CHARACTERISTICS:
LENGTH: 197 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
FRAGMENT TYPE: N-terminal
US-08-456-461-27

Query Match 20.4%; Score 153.5; DB 2; Length 197;
Best Local Similarity 30.4%; Pred. No. 2.5e-09;
Matches 42; Conservative 21; Mismatches 56; Indels 19; Gaps 8;
QY 3 SCGSPPTILGRISYSTPIAVGIVIRYSCGIFELIGKSLLC-ITKDKVGTWDRKAP 61
DB 63 SCRNPDVNGMV-HVIRKIGFGSOIKYSCIKYKGLIGSSSAPGIIISGDIV--IWDNETP 119
QY 62 KCEYFNKYSSCEPIVPGYKIRGISTYRHGDSVTFACKTN-----FSMNGKNSVMCOA 115
DB 120 ICDRIPT--CGLPPTITNGFISTNPFNFHYGSVYPTNFGSGGKVELVGEPSVCTS 177
QY 116 NN---MW-GPTRLPTCV 128
DB 178 NDDQVCWISGPA--PQCI 193

RESULT 6
US-08-769-967A-27
Sequence 27, Application US/08769967A
Patent No. 5859223
GENERAL INFORMATION:
APPLICANT: Mossakowska, Barbara E.L.
APPLICANT: Smith, Richard A.G.
APPLICANT: Dodd, Ian
APPLICANT: Freeman, Anne Mary
TITLE OF INVENTION: Soluble (R1) Derivatives
NUMBER OF SEQUENCES: 33
CORRESPONDENCE ADDRESS:
ADDRESSEE: SmithKline Beecham Corporate Intellectual Property

STREET: P.O. Box 1539
CITY: King of Prussia
STATE: Pennsylvania
COUNTRY: USA
ZIP: 19406
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/769,967A
FILING DATE:
CLASSIFICATION: 536
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/440,564
FILING DATE: 15-May-1995
ATTORNEY/AGENT INFORMATION:
NAME: King, William T.
REGISTRATION NUMBER: 30,954
REFERENCE/DOCKET NUMBER: P30423C.
TELECOMMUNICATION INFORMATION:
TELEPHONE: (610) 270-5364
TELEFAX: (610) 270-5090
INFORMATION FOR SEQ ID NO: 27:
SEQUENCE CHARACTERISTICS:
LENGTH: 197 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
FRAGMENT TYPE: N-terminal
US-08-769-967A-27

Query Match 20.4%; Score 153.5; DB 2; Length 197;
Best Local Similarity 30.4%; Pred. No. 2.5e-09;
Matches 42; Conservative 21; Mismatches 56; Indels 19; Gaps 8;
QY 3 SCGSPPTILGRISYSTPIAVGIVIRYSCGIFELIGKSLLC-ITKDKVGTWDRKAP 61
DB 63 SCRNPDVNGMV-HVIRKIGFGSOIKYSCIKYKGLIGSSSAPGIIISGDIV--IWDNETP 119
QY 62 KCEYFNKYSSCEPIVPGYKIRGISTYRHGDSVTFACKTN-----FSMNGKNSVMCOA 115
DB 120 ICDRIPT--CGLPPTITNGFISTNPFNFHYGSVYPTNFGSGGKVELVGEPSVCTS 177
QY 116 NN---MW-GPTRLPTCV 128
DB 178 NDDQVCWISGPA--PQCI 193

RESULT 7
US-08-824-692-3;
Sequence 32, Application US/08824692
Patent No. 60,7703
GENERAL INFORMATION:
APPLICANT: Kinders, Robert J.
APPLICANT: Enfield, David L.
APPLICANT: Bass, G. Michael
TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR SCREENING
OPERATING SYSTEM: FOR OR MODULATING A TUMOR ASSOCIATED ANTI-EN
NUMBER OF SEQUENCES: 38
CORRESPONDENCE ADDRESS:
ADDRESSEE: SEED and HERRY LLP
STREET: 6-00 Columbia Center, 701 - 11th Avenue
CITY: Seattle
STATE: Washington
COUNTRY: USA
ZIP: 98104
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.40

DB 198 QKUSTGCKGPPPPHNDITSPFSAVAASSVYQCONLYOLEGNKRKTOR-NCOWSR 256
QY 121 PTRLPTCV 128
DB 257 PPKLHPCV 265

RESULT 40
US-08-296-014A-4
PATENT NO. 5514582
APPLICANT: CAPON, DANIEL J.; LASKY, LAURENCE A.
TITLE OF INVENTION: RECOMBINANT DNA ENCODING HYBRID
IMMUNOGLOBULINS
NUMBER OF SEQUENCES: 4
CURRENT APPLICATION DATA:
APPLICATION NUMBER: 05/08/185,670
FILING DATE: 21 JAN 1994
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: 986,941
FILING DATE: 08 DEC 1992
APPLICATION NUMBER: 808,122
FILING DATE: 16 DEC 1993
APPLICATION NUMBER: 440,625
FILING DATE: 22 NOV 1989
APPLICATION NUMBER: 315,015
FILING DATE: 23 FEB 1989
SEQ ID NO: 4
LENGTH: 126
5514582/4

Query Match
Best Local Similarity 32.4%; Pred. No. 1.5e 08,
Matches 41; Conservative 13; Mismatches 59; Indels 14; Gaps 5;

QY 4 CGSPPPLNGRISYSYIPIAVGVIRYSGTFRLLIGESKSLICITKDKVIGTWKPAKPC 63
DB 5 CPSRDRDNGDLDGG- VIFGSIYSYSCNSGYHLIGESKSYGELSGTGMWNNPRAPIC 62
QY 64 EYENKYSPPPIVPG---CYKRWGTPPYPIPISTVFAKTNFSMNGKSVWQANNMWG 120
DB 63 ESKVKUS--PPISLNGRHNYE---DFTYDGSVVIYSCNSGYSLICNSVLCSSGGHWS 116
QY 121 PTRLPTCV 127
DB 117 P---PTC 120

RESULT 41
US-08-296-014A-4
PATENT NO. 5716844
GENERAL INFORMATION:
APPLICANT: Ho, Bow
TITLE OF INVENTION: The Cloned Factor C cDNA of the
TITLE OF INVENTION: Singapore Horseshoe Crab, Carcinoscorpius
NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESS:
ADDRESSEE: Hirsch, Stewart, Kolasch & Hirsch
STREET: 8110 Gatehouse Road, Suite 500 East
CITY: Falls Church
STATE: Virginia
COUNTRY: USA
ZIP: 22042
COMPUTER READABLE FORM:
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/296,014A

FILING DATE: 435
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Murphy, Jr., Gerald M.
REGISTRATION NUMBER: 28,977
REFERENCE/DOCKET NUMBER: 1781-105
TELEPHONE: (703) 205-8000
TELEFAX: (703) 205-8050
TELEX: 248345
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 3019 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-296-014A-4

Query Match
Best Local Similarity 19.2%; Score 144; DB 1; Length 1019;
Matches 38; Conservative 31; Mismatches 50; Indels 46; Gaps 8;

QY 2 ISC-----GSPPLI-----NGRISYSYIPIAVGVIRYSGTFRLLIGESKSL 44
DB 130 ISCLPNCQWSNFPKCTIRCAVSSSHGKVNALSDMIETATIRFSQSHVYVLIQQH 239
QY 45 LCIILKLVIGTWKPAKPCYEYENKYSPP-ELVPGGKRLK-----GSTPYHRLSV 95
DB 240 TC-----QGNQWNGQIPQCK---NIVFGTDLDFNIAELKVKLVWQYKRGFP- 290
QY 96 IFACKINF-SMNGKSVWQANNMWGPIPIPICTCV 130
DB 291 TVTCSNYELMGEDTLKCNPDGWSGSGQ-PSCKV 324

RESULT 42
US-08-596-405-4
PATENT NO. 5858706
GENERAL INFORMATION:
APPLICANT: Ho, Bow
TITLE OF INVENTION: The Cloned Factor C cDNA of the
TITLE OF INVENTION: Singapore Horseshoe Crab, Carcinoscorpius
NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESS:
ADDRESSEE: Hirsch, Stewart, Kolasch & Hirsch
STREET: 8110 Gatehouse Road, Suite 500 East
CITY: Falls Church
STATE: Virginia
COUNTRY: USA
ZIP: 22042
COMPUTER READABLE FORM:
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/596,405
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Murphy, Jr., Gerald M.
REGISTRATION NUMBER: 28,977
REFERENCE/DOCKET NUMBER: 1781-105
TELEPHONE: (703) 205-8000
TELEFAX: (703) 205-8050
TELEX: 248345
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 3019 amino acids

TYPE: amino acid
topology: linear
MOLECULE TYPE: protein

US 08 296 014A 4

Query Match 19.2% Score 144; DB 2; Length 1019;
Best Local Similarity 24.5%; Pred. No. 2.4e-07;
Matches 48; Conservative 41; Mismatches 50; Indels 46; Gaps 8;

QY 2 150 - GSEPTIL - NGRISYSTTAVGVTKVNSGTFRLGKSL 44
DB 180 ISCLPNCQWSNFPKCIREFCAWVSSPEIKKVNALSGMIEGATLRFSDSYVYLLIGQETL 249
QY 45 LQILKIKVIGWIKPAPKTEFNKYSSP EFIVGGYKIR - GSPVPHHDSV 95
DB 240 TC - - - - - QNGNQGQIPQCK - - - - - NIVECPHLLPWNHAEKVKVGVVQVYQGF GQTEV 290
QY 96 TEACKNFSGNKNKSVWQANNMKGPTRLPTCVSV 140
DB 291 TYLCSNFTLMGPDILKNDPSWSGSO PSQVAV 424

RESULT 44

Sequence 4; Application US/08/014A 2
Patent No. 5716834

GENERAL INFORMATION:

APPLICANT: Ding, Jack Ling
TITLE OF INVENTION: The Cloned Factor C cDNA of the
TITLE OF INVENTION: Singapore Horseshoe Crab, Carcinus scaber
TITLE OF INVENTION: rotundicauda and purification of Factor C Proenzyme
NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESS:
ADDRESSEE: Birch, Stewart, Kolasch & Birch
STREET: 8110 Gatehouse Road, Suite 500 East
CITY: Falls Church
STATE: Virginia
COUNTRY: USA
ZIP: 22042

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC DOS/MS DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/014A 2
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION NUMBER: 08/096,405
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Murphy, Jr., Gerald M.
REGISTRATION NUMBER: 28,977
REFERENCE/TWCKET NUMBER: 1781-105P
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 205-8000
TELEFAX: (703) 205-8050
HIEEX: 248445
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 1019 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US 08 296 014A 2

Query Match

Best Local Similarity 24.5%; Score 144; DB 2; Length 1019;
Matches 48; Conservative 41; Mismatches 50; Indels 46; Gaps 8;

QY 2 150 - GSEPTIL - NGRISYSTTAVGVTKVNSGTFRLGKSL 44
DB 244 ISCLPNCQWSNFPKCIREFCAWVSSPEIKKVNALSGMIEGATLRFSDSYVYLLIGQETL 404
QY 45 LQILKIKVIGWIKPAPKTEFNKYSSP EFIVGGYKIR - GSPVPHHDSV 95
DB 304 TC - - - - - QNGNQGQIPQCK - - - - - NIVECPHLLPWNHAEKVKVGVVQVYQGF GQTEV 404
QY 96 TEACKNFSGNKNKSVWQANNMKGPTRLPTCVSV 140
DB 355 TYLCSNFTLMGPDILKNDPSWSGSO PSQVAV 488

RESULT 45

Query Match 19.2% Score 144; DB 2; Length 1019;
Best Local Similarity 24.5%; Pred. No. 2.4e-07;
Matches 48; Conservative 41; Mismatches 50; Indels 46; Gaps 8;

QY 2 150 - GSEPTIL - NGRISYSTTAVGVTKVNSGTFRLGKSL 44

DB 180 ISCLPNCQWSNFPKCIREFCAWVSSPEIKKVNALSGMIEGATLRFSDSYVYLLIGQETL 249
QY 45 LQILKIKVIGWIKPAPKTEFNKYSSP EFIVGGYKIR - GSPVPHHDSV 95
DB 240 TC - - - - - QNGNQGQIPQCK - - - - - NIVECPHLLPWNHAEKVKVGVVQVYQGF GQTEV 290
QY 96 TEACKNFSGNKNKSVWQANNMKGPTRLPTCVSV 140
DB 291 TYLCSNFTLMGPDILKNDPSWSGSO PSQVAV 424

RESULT 44

Sequence 2; Application US/08296014A
Patent No. 5716834

GENERAL INFORMATION:

APPLICANT: Ding, Jack Ling
TITLE OF INVENTION: The Cloned Factor C cDNA of the
TITLE OF INVENTION: Singapore Horseshoe Crab, Carcinus scaber
TITLE OF INVENTION: rotundicauda and purification of Factor C Proenzyme
NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESS:
ADDRESSEE: Birch, Stewart, Kolasch & Birch
STREET: 8110 Gatehouse Road, Suite 500 East
CITY: Falls Church
STATE: Virginia
COUNTRY: USA
ZIP: 22042

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC DOS/MS DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/296,014A
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Murphy, Jr., Gerald M.
REGISTRATION NUMBER: 28,977
REFERENCE/TWCKET NUMBER: 1781-105P
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 205-8000
TELEFAX: (703) 205-8050
HIEEX: 248445
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 1083 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US 08-296-014A-2

Query Match

Best Local Similarity 24.5%; Score 144; DB 1; Length 1083;
Matches 48; Conservative 41; Mismatches 50; Indels 46; Gaps 8;

QY 2 150 - GSEPTIL - NGRISYSTTAVGVTKVNSGTFRLGKSL 44
DB 244 ISCLPNCQWSNFPKCIREFCAWVSSPEIKKVNALSGMIEGATLRFSDSYVYLLIGQETL 404
QY 45 LQILKIKVIGWIKPAPKTEFNKYSSP EFIVGGYKIR - GSPVPHHDSV 95
DB 304 TC - - - - - QNGNQGQIPQCK - - - - - NIVECPHLLPWNHAEKVKVGVVQVYQGF GQTEV 404
QY 96 TEACKNFSGNKNKSVWQANNMKGPTRLPTCVSV 140
DB 355 TYLCSNFTLMGPDILKNDPSWSGSO PSQVAV 488

RESULT 45

```

US 08 596 405 2
? Sequence 2: Application US/08596405
? Patent No. 5858706
? GENERAL INFORMATION:
? APPLICANT: Ding, Jack Ling
? APPLICANT: Ho, How
? TITLE OF INVENTION: The Cloned Factor C cDNA of the
? TITLE OF INVENTION: Singapore Herpesheo Crab, Carcinoscorpius
? TITLE OF INVENTION: rotundicauda and Purification of Factor C Proenzyme
? NUMBER OF SEQUENCES: 4
? CORRESPONDENCE ADDRESS:
? ADDRESSEE: Birch, Stewart, Kolasch & Birch
? STREET: 8110 Gatehouse Road, Suite 500 East
? CITY: Falls Church
? STATE: Virginia
? COUNTRY: USA
? ZIP: 22042
? COMPUTER READABLE FORM:
? MEDIUM TYPE: Floppy disk
? COMPUTER: IBM PC compatible
? OPERATING SYSTEM: PC DOS/MS DOS
? SOFTWARE: Patent In Release #1.0, Version #1.25
? CURRENT APPLICATION DATA:
? APPLICATION NUMBER: US/08/596,405
? FILING DATE:
? CLASSIFICATION: 435
? ATTORNEY/AGENT INFORMATION:
? NAME: Murphy, J. Gerald M.
? REGISTRATION NUMBER: 28,977
? REFERENCE/DOCKET NUMBER: 1781 1050
? TELECOMMUNICATION INFORMATION:
? TELEPHONE: (703) 205-8000
? TELEFAX: (703) 205-8050
? TELEX: 248345
? INFORMATION FOR SEQ ID NO: 2:
? SEQUENCE CHARACTERISTICS:
? LENGTH: 1083 amino acids
? TYPE: amine acid
? TOPOLOGY: Linear
? MOLECULE TYPE: protein
US 08 596 405 2

Query Match 19.2%; Score 144; DB 2; Length 1083;
Best Local Similarity 24.5%; Pred. No. 2.5e-07;
Matches 38; Conservative 41; Mismatches 50; Indels 36; Gaps 8;

QY 2 1SC--- -GSPPPIL --- NGRISYKSTPIAVGTIVIRYSGSTFRLIGEKSL 44
   ||| | | | | | | | | | | | | | | | | | | | | | | | | |
DB 244 1SCLNCGWSNFPKCIKECAWSSPEKGVNAISGDMLEGATLRFSCDSPPYLLIGQFTL 303
   | | | | | | | | | | | | | | | | | | | | | | | | | |

QY 45 1G1TKKVGWTKVAPKCYENKYSSP--EPVPGGYKTR-----GSTPYRHGDSV 95
   | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 304 TC --QNGGWNKGIPOCK NLVECPDLTPVNIIEHKVIGVEQYKQGP--QGTEV 354
   | | | | | | | | | | | | | | | | | | | | | | | | | |

QY 96 1EAKKINFSGNCKSVWCQANNMGPIKPLICVSV 140
   | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 355 TYTSGNIFLMGFDTLKNFDSNSGSG PSCVKV 388
   | | | | | | | | | | | | | | | | | | | | | | | | | |

```

Search completed: November 6, 2002, 16:06:53
Job time : 13.461 secs



Genome version 5.1.3
Copyright (c) 1994-2002 Computer Ltd.

OM protein protein search, using sw model

Run on: November 6, 2002, 16:04:59 : Search time 26.8996 Seconds
(without alignments)
553,312 Million cell updates/sec

Title: US 09 834-309 4
RefSeq score: 751
Sequence: 1 GILSGSPPLILNRISSYST ANNMWSPPTPTVSVPEPLE 134

Scoring table: BLASTM62
Gap: 10.0, Gapext 0.5

Searched: 747574 seqs, 11107396 residues
Total number of hits satisfying chosen parameters: 747574

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post processing: Minimum Match ok
Maximum Match 100%
Listing first 100 summaries

Database :			
A_Geneseq_032802.8			
1:	/SIDSI/qqdta/q/geneseq/geneseq-emb1/AA1980.DAT:	12	216
2:	/SIDSI/qqdta/q/geneseq/geneseq-emb1/AA1981.DAT:	13	215
3:	/SIDSI/qqdta/q/geneseq/geneseq-emb1/AA1982.DAT:	14	215
4:	/SIDSI/qqdta/q/geneseq/geneseq-emb1/AA1983.DAT:	15	215
5:	/SIDSI/qqdta/q/geneseq/geneseq-emb1/AA1984.DAT:	16	210
6:	/SIDSI/qqdta/q/geneseq/geneseq-emb1/AA1985.DAT:	17	210
7:	/SIDSI/qqdta/q/geneseq/geneseq-emb1/AA1986.DAT:	18	204
8:	/SIDSI/qqdta/q/geneseq/geneseq-emb1/AA1987.DAT:	19	203
9:	/SIDSI/qqdta/q/geneseq/geneseq-emb1/AA1988.DAT:	20	203
10:	/SIDSI/qqdta/q/geneseq/geneseq-emb1/AA1989.DAT:	21	203
11:	/SIDSI/qqdta/q/geneseq/geneseq-emb1/AA1990.DAT:	22	203
12:	/SIDSI/qqdta/q/geneseq/geneseq-emb1/AA1991.DAT:	23	203
13:	/SIDSI/qqdta/q/geneseq/geneseq-emb1/AA1992.DAT:	24	203
14:	/SIDSI/qqdta/q/geneseq/geneseq-emb1/AA1993.DAT:	25	203
15:	/SIDSI/qqdta/q/geneseq/geneseq-emb1/AA1994.DAT:	26	203
16:	/SIDSI/qqdta/q/geneseq/geneseq-emb1/AA1995.DAT:	27	203
17:	/SIDSI/qqdta/q/geneseq/geneseq-emb1/AA1996.DAT:	28	203
18:	/SIDSI/qqdta/q/geneseq/geneseq-emb1/AA1997.DAT:	29	203
19:	/SIDSI/qqdta/q/geneseq/geneseq-emb1/AA1998.DAT:	30	203
20:	/SIDSI/qqdta/q/geneseq/geneseq-emb1/AA1999.DAT:	31	203
21:	/SIDSI/qqdta/q/geneseq/geneseq-emb1/AA2000.DAT:	32	203
22:	/SIDSI/qqdta/q/geneseq/geneseq-emb1/AA2001.DAT:	33	203

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES			
Result No.	Score	Query Match Length DB ID	Description
1	751	100.0	1087 12 AA11139 H lymphocyte membr
2	219	29.2	1540 12 AA11982 Partial human comp
3	219	29.2	1929 22 AP00103 Novel human diagno
4	219	29.2	1940 19 AA45894 Human complement c
5	219	29.2	1949 12 AA11810 Human complement c
6	219	29.2	2039 20 AA55751 Human C3b/C4b rec
7	219	29.2	2044 22 AP00287 Novel human diagno
8	219	29.2	2044 22 AP01782 Human CR1 protein
9	219	29.2	2044 22 AA49224 Human polypeptide
10	219	29.2	2044 22 AA41010 Human polypeptide
11	219	29.2	2317 19 AA19219 CR1 protein, Homo

216	28.8	263	22	AA4884	Newcastle virus Italia
215	28.6	263	10	AA192004	Dodged megaloc
215	28.6	263	20	AA12981	Vac. virus comp. lomen
215	28.6	263	21	AA11304	Complement inhibi
212	28.2	2039	14	AA16763	CR1, Homo sapiens
211	28.1	450	20	AA15755	Human CR1 protein
210	28.0	263	20	AA12983	Vaccinia virus SPI
210	28.0	263	20	AA15986	Mutated VEP avian
204	27.2	646	20	AA15578	Human CR1 protein
203	27.1	579	19	AA19924	Amino acid sequen
199	26.5	438	20	AA15576	Human CR1 protein
199	26.5	450	20	AA15573	Human CR1 protein
199	26.5	496	20	AA15572	Human CR1 protein
199	26.5	581	12	AA11340	Human CR1 protein
199	26.5	778	19	AA17317	Human CR1 protein
198	26.4	133	15	AA17156	Amino acid sequen
198	26.4	254	15	AA17154	Sequence of solubl
198	26.4	450	20	AA15574	Human CR1 protein
198	26.4	453	20	AA15577	Human CR1 protein
196	26.1	481	13	AA12902	CR1-4 (Gmbic. acids
196	26.1	483	13	AA12901	CR1-4 (Gmbic. acids
196	26.1	543	13	AA12854	CR1-4 (45E, 47Y) a
196	26.1	543	13	AA12854	CR1-4 (45E) analo
196	26.1	543	13	AA12854	CR1-4 (47Y) analo
196	26.1	543	13	AA12854	CR1-4 (44T, 47U, 4
196	26.1	543	13	AA12854	CR1-4 (42S, 44S, 5
196	26.1	543	13	AA12854	CR1-4 (47V, 49K) a
196	26.1	543	13	AA12854	CR1-4 (44K, 65I) a
196	26.1	543	13	AA12855	CR1-4 (44K) analo
196	26.1	543	13	AA12855	CR1-4 (45T) analo
196	26.1	543	13	AA12855	CR1-4 (48T, 79D) a
196	26.1	543	13	AA12855	CR1-4 (48K, 87N) a
196	26.1	543	13	AA12855	CR1-4 (44T, 47U, 4
196	26.1	543	13	AA12855	CR1-4 (42S, 44S, 5
196	26.1	543	13	AA12855	CR1-4 (47V, 49K) a
196	26.1	543	13	AA12855	CR1-4 (44K, 65I) a
196	26.1	543	13	AA12855	CR1-4 (44K) analo
196	26.1	543	13	AA12855	CR1-4 (45T) analo
196	26.1	543	13	AA12855	CR1-4 (48T, 79D) a
196	26.1	543	13	AA12855	CR1-4 (48K, 87N) a
196	26.1	543	13	AA12855	CR1-4 (44T, 47U, 4
196	26.1	543	13	AA12855	CR1-4 (42S, 44S, 5
196	26.1	543	13	AA12855	CR1-4 (47V, 49K) a
196	26.1	543	13	AA12855	CR1-4 (44K, 65I) a
196	26.1	543	13	AA12855	CR1-4 (44K) analo
196	26.1	543	13	AA12855	CR1-4 (45T) analo
196	26.1	543	13	AA12855	CR1-4 (48T, 79D) a
196	26.1	543	13	AA12855	CR1-4 (48K, 87N) a
196	26.1	543	13	AA12855	CR1-4 (44T, 47U, 4
196	26.1	543	13	AA12855	CR1-4 (42S, 44S, 5
196	26.1	543	13	AA12855	CR1-4 (47V, 49K) a
196	26.1	543	13	AA12855	CR1-4 (44K, 65I) a
196	26.1	543	13	AA12855	CR1-4 (44K) analo
196	26.1	543	13	AA12855	CR1-4 (45T) analo
196	26.1	543	13	AA12855	CR1-4 (48T, 79D) a
196	26.1	543	13	AA12855	CR1-4 (48K, 87N) a
196	26.1	543	13	AA12855	CR1-4 (44T, 47U, 4
196	26.1	543	13	AA12855	CR1-4 (42S, 44S, 5
196	26.1	543	13	AA12855	CR1-4 (47V, 49K) a
196	26.1	543	13	AA12855	CR1-4 (44K, 65I) a
196	26.1	543	13	AA12855	CR1-4 (44K) analo
196	26.1	543	13	AA12855	CR1-4 (45T) analo
196	26.1	543	13	AA12855	CR1-4 (48T, 79D) a
196	26.1	543	13	AA12855	CR1-4 (48K, 87N) a
196	26.1	543	13	AA12855	CR1-4 (44T, 47U, 4
196	26.1	543	13	AA12855	CR1-4 (42S, 44S, 5
196	26.1	543	13	AA12855	CR1-4 (47V, 49K) a
196	26.1	543	13	AA12855	CR1-4 (44K, 65I) a
196	26.1	543	13	AA12855	CR1-4 (44K) analo
196	26.1	543	13	AA12855	CR1-4 (45T) analo
196	26.1	543	13	AA12855	CR1-4 (48T, 79D) a
196	26.1	543	13	AA12855	CR1-4 (48K, 87N) a
196	26.1	543	13	AA12855	CR1-4 (44T, 47U, 4
196	26.1	543	13	AA12855	CR1-4 (42S, 44S, 5
196	26.1	543	13	AA12855	CR1-4 (47V, 49K) a
196	26.1	543	13	AA12855	CR1-4 (44K, 65I) a
196	26.1	543	13	AA12855	CR1-4 (44K) analo
196	26.1	543	13	AA12855	CR1-4 (45T) analo
196	26.1	543	13	AA12855	CR1-4 (48T, 79D) a
196	26.1	543	13	AA12855	CR1-4 (48K, 87N) a
196	26.1	543	13	AA12855	CR1-4 (44T, 47U, 4
196	26.1	543	13	AA12855	CR1-4 (42S, 44S, 5
196	26.1	543	13	AA12855	CR1-4 (47V, 49K) a
196	26.1	543	13	AA12855	CR1-4 (44K, 65I) a
196	26.1	543	13	AA12855	CR1-4 (44K) analo
196	26.1	543	13	AA12855	CR1-4 (45T) analo
196	26.1	543	13	AA12855	CR1-4 (48T, 79D) a
196	26.1	543	13	AA12855	CR1-4 (48K, 87N) a
196	26.1	543	13	AA12855	CR1-4 (44T, 47U, 4
196	26.1	543	13	AA12855	CR1-4 (42S, 44S, 5
196	26.1	543	13	AA12855	CR1-4 (47V, 49K) a
196	26.1	543	13	AA12855	CR1-4 (44K, 65I) a
196	26.1	543	13	AA12855	CR1-4 (44K) analo
196	26.1	543	13	AA12855	CR1-4 (45T) analo
196	26.1	543	13	AA12855	CR1-4 (48T, 79D) a
196	26.1	543	13	AA12855	CR1-4 (48K, 87N) a
196	26.1	543	13	AA12855	CR1-4 (44T, 47U, 4
196	26.1	543	13	AA12855	CR1-4 (42S, 44S, 5
196	26.1	543	13	AA12855	CR1-4 (47V, 49K) a
196	26.1	543	13	AA12855	CR1-4 (44K, 65I) a
196	26.1	543	13	AA12855	CR1-4 (44K) analo
196	26.1	543	13	AA12855	CR1-4 (45T) analo
196	26.1	543	13	AA12855	CR1-4 (48T, 79D) a
196	26.1	543	13	AA12855	CR1-4 (48K, 87N) a
196	26.1	543	13	AA12855	CR1-4 (44T, 47U, 4
196	26.1	543	13	AA12855	CR1-4 (42S, 44S, 5
196	26.1	543	13	AA12855	CR1-4 (47V, 49K) a
196	26.1	543	13	AA12855	CR1-4 (44K, 65I) a
196	26.1	543	13	AA12855	CR1-4 (44K) analo
196	26.1	543	13	AA12855	CR1-4 (45T) analo
196	26.1	543	13	AA12855	CR1-4 (48T, 79D) a
196	26.1	543	13	AA12855	CR1-4 (48K, 87N) a
196	26.1	543	13	AA12855	CR1-4 (44T, 47U, 4
196	26.1	543	13	AA12855	CR1-4 (42S, 44S, 5
196	26.1	543	13	AA12855	CR1-4 (47V, 49K) a
196	26.1	543	13	AA12855	CR1-4 (44K, 65I) a
196	26.1	543	13	AA12855	CR1-4 (44K) analo
196	26.1	543	13	AA12855	CR1-4 (45T) analo
196	26.1	543	13	AA12855	CR1-4 (48T, 79D) a
196	26.1	543	13	AA12855	CR1-4 (48K, 87N) a
196	26.1	543	13	AA12855	CR1-4 (44T, 47U, 4
196	26.1	543	13	AA12855	CR1-4 (42S, 44S, 5
196	26.1	543	13	AA12855	CR1-4 (47V, 49K) a
196	26.1	543	13	AA12855	CR1-4 (44K, 65I) a
196	26.1	543	13	AA12855	CR1-4 (44K) analo
196	26.1	543	13	AA12855	CR1-4 (45T) analo
196	26.1	543	13	AA12855	CR1-4 (48T, 79D) a
196	26.1	543	13	AA12855	CR1-4 (48K, 87N) a
196	26.1	543	13	AA12855	CR1-4 (44T, 47U, 4
196	26.1	543	13	AA12855	CR1-4 (42S, 44S, 5
196	26.1	543	13	AA12855	CR1-4 (47V, 49K) a
196	26.1	543	13	AA12855	CR1-4 (44K, 65I) a
196	26.1	543	13	AA12855	CR1-4 (44K) analo
196	26.1	543	13	AA12855	CR1-4 (45T) analo
196	26.1	543	13	AA12855	CR1-4 (48T, 79D) a
196	26.1	543	13	AA12855	CR1-4 (48K, 87N) a
196	26.1	543	13	AA12855	CR1-4 (44T, 47U, 4
196	26.1	543	13	AA12855	CR1-4 (42S, 44S, 5
196	26.1	543	13	AA12855	CR1-4 (47V, 49K) a
196	26.1	543	13	AA12855	CR1-4 (44K, 65I) a
196	26.1	543	13	AA12855	CR1-4 (44K) analo
196	26.1	543	13	AA12855	CR1-4 (45T) analo
196	26.1	543	13	AA12855	CR1-

XX 26 SEP 1989; HQUS 0412745.
 XX 26 SEP 1990; HQUS 0912449.
 XX (CCEL) T CELL SCI INC.
 PA (CJJO) JOHNS HOPKINS UNIVERSITY.
 PA (HRC) BRIGHAM AND WOMEN'S HOSPITAL.
 XX Pearson DT, Klickstein LB, Wong WW, Carson GR, Boh M, Concino MF;
 PI Makrides SC, Marsh RC;
 XX WPI: 2001-63362/73.
 DR N-PSDB: AAQ11643.
 PS Human complement receptor type 1 gene, encoded proteins and
 PI fragments for treatment of immune disorders, myocardial infarct,
 PI damage due to inflammation and in treatment of thrombosis
 XX Disclosure: Fig 5; 24pp; English.
 XX This sequence comprises three of the four tandem, direct, long
 CC homologous repeats of the full-length F allozyme of C3b/C4b binding domain,
 CC absent. Each LIR might represent a single C3b/C4b binding domain,
 CC making the receptor multivalent. The LIRs are composed of 7 short
 CC consensus repeats of 60-70 residues resembling the SCR's of other
 CC C3/C4 binding proteins. The protein and fragments of it having C3b
 CC and/or C4b binding activity can be used to treat immune disorders
 CC or disorders involving inappropriate complement activity.
 CC See also AAQ11642.
 XX Sequence 1547 AA;
 SQ Query Match 29.2%; Score 219; DB 12; Length 1537;
 Best Local Similarity 36.7%; Pred. No. 3.7e-13;
 Matches 51; Conservative 17; Mismatches 55; Indels 16; Gaps 6;
 QY 2 ISGSHPPHNGKISYVS---TPVAGVIVVYSC-----SGTFRIGKSLCTTKDKV 52
 DB 1015 ISCEPPTISNG--DFYNNPTSFHNTVVYQHTGPKDEQLFELVGFSTVYTSKDDQ 1072
 QY 53 DGTWDKPAKCEYFNKYSSCPPIVPGGYEIKGISTP-YRIGDSVTFACKTNFSGMGNKSV 111
 DB 1073 VGVWSSPPPHCTSTNKK---CTAFVEVENAIVPNRSPFSSTETTRPRQPGFVWVGSHTV 1129
 QY 112 WQANNMW:PTRLPTCVSV 130
 DB 1130 QQTNGRWGP-KLPHCGRV 1147
 RESULT 4
 ID AAG00104
 XX AAG00104 standard; protein; 1929 AA.
 XX AAG00104;
 XX 14 FEB 2002 (first entry)
 XX Novel human diagnostic protein #94.
 XX Human; chromosome mapping; gene mapping; gene therapy; forensic;
 XX food supplement; medical imaging; diagnostic; genetic disorder.
 XX Homo sapiens.
 XX W0200175067 A2.
 XX 11 OCT 2001.
 XX 30 MAR 2001; 2001WO/0508631.
 XX 31 MAR 2000; 2000US-0540217.
 XX 24 AUG 2000; 2000US-0649767.
 XX

PA (HYSE-) HYSEQ INC.
 PI Drmanac RT, Liu C, Tanq YT;
 XX WPI: 2001-63362/73.
 DR N-PSDB: AAS64290.
 XX New isolated polynucleotide and encoded polypeptides, useful in
 PT diagnostics, forensics, gene mapping, identification of mutations
 PT responsible for genetic disorders or other traits and to assess
 PT bioactivity
 XX Claim 20; SEQ ID No 30462; 103pp; English.
 XX The invention relates to isolated polynucleotide (I) and
 CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
 CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
 CC and gene mapping, and in recombinant production of (II). The
 CC polynucleotides are also used in diagnostics as expressed sequence tags
 CC for identifying expressed genes. (I) is useful in gene therapy techniques
 CC to restore normal activity of (II) or to treat disease states involving
 CC (II). (II) is useful for generating antibodies against it, detection of
 CC quantitation a polypeptide in tissue, at molecular weight markers and as
 CC a food supplement. (II) and its binding partners are useful for medical
 CC imaging of sites expressing (II). (I) and (II) are useful for treating
 CC disorders involving aberrant protein expression or biological activity.
 CC The polypeptide and polynucleotide sequences have applications in
 CC diagnostics, forensics, gene mapping, identification of mutations
 CC responsible for genetic disorders or other traits to assess bioactivity
 CC and to produce other types of data and products dependent on DNA and
 CC amino acid sequences. AAG00104-AAG30477 represent novel human
 CC diagnostic amino acid sequences of the invention.
 CC Note: The sequence data for this patent did not appear in the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at http://wipo.int/pub/published_pat_sequences.
 XX Sequence 1929 AA;
 SQ Query Match 29.2%; Score 219; DB 22; Length 1929;
 Best Local Similarity 36.7%; Pred. No. 4.8e-14;
 Matches 51; Conservative 17; Mismatches 55; Indels 16; Gaps 6;
 QY 2 ISGSHPPHNGKISYVS---TPVAGVIVVYSC-----SGTFRIGKSLCTTKDKV 52
 DB 1407 ISCEPPTISNG--DFYNNPTSFHNTVVYQHTGPKDEQLFELVGFSTVYTSKDDQ 1464
 QY 53 DGTWDKPAKCEYFNKYSSCPPIVPGGYEIKGISTP-YRIGDSVTFACKTNFSGMGNKSV 111
 DB 1465 VGVWSSPPPHCTSTNKK---CTAFVEVENAIVPNRSPFSSTETTRPRQPGFVWVGSHTV 1521
 QY 112 WQANNMW:PTRLPTCVSV 130
 DB 1522 QQTNGRWGP-KLPHCGRV 1539
 RESULT 4
 ID AAW45899
 XX AAW45899 standard; peptide; 1930 AA
 XX AAW45899;
 XX 30-JUN-1998 (first entry)
 XX Human complement receptor 1 (residues 1-1924).
 XX Membrane binding element; thrombotic disease; soluble proteins;
 XX complement related disease; integral membrane proteins; inflammation;
 XX short consensus repeat; SCR 1-3; C3b; complement receptor type 1.
 XX Homo sapiens.
 XX Key location/Qualifiers
 XX Cross-links 1930

XX Homo sapiens.
 XX OS 055981481 A.
 XX ID 09 NOV 1999.
 XX AC 9508-0470652.
 XX PR 03 APR 1989; 8908-032865.
 XX PR 06 DEC 1974; 7408-0450248.
 XX PR 24 FEB 1983; 8408-0026144.
 XX PR 01 APR 1988; 8808-0176542.
 XX PA (YCHO) UNIV JOHNS HOPKINS.
 XX PR (BCHM) BRIGHAM & WOMEN'S HOSPITAL.
 XX PR (AVAN) AVANT IMMUNOTHERAPEUTICS INC.
 XX PR Connelley MF, Wood WW, Makrides SN, Klickstein LB, Fearon DF, Ip SH, Marsh HC, Carlson GR.
 XX PR WP: 1999, 643351/54.
 XX PR N PSDB; AAZ40150.
 XX PR A human C3b/C4b receptor (C3R) protein having antiinflammatory and
 XX PR cardiant activity
 XX PR Disclosure; Fig 1A-P; 87pp; English.
 XX CC The invention relates to a human C3b/C4b receptor (C3R) protein, the C3R
 XX CC protein or fragment is expressed as a cell-surface protein on the surface
 XX CC of a non-human cell and exhibits a complement regulatory activity of full
 XX CC length human C3R as expressed on erythrocytes. The C3R function in vivo
 XX CC may be mediated through the inhibition of complement pathway enzymes. The
 XX CC soluble C3R protein exhibits a complement regulatory activity, and this
 XX CC may be used to prevent reperfusion injury, inhibit Arthus reaction, and
 XX CC neutrophil mediated tissue damage, and reduce myocardial infarct size,
 XX CC and inflammation. The C3R protein and its fragments can also be used in
 XX CC the treatment of conditions which involve unwanted complement activity,
 XX CC e.g. shock lung, tissue damage due to burn, or ischemic heart conditions,
 XX CC and autoimmune disorders. C3R proteins, analogues, derivatives, and anti
 XX CC C3R antibodies are used in assays, and diagnostics. The present sequence
 XX CC represents the human C3R protein.
 XX SQ Sequence 2039 AA;
 Query Match 29.2%; Score 219; DB 20; Length 2039;
 Best Local Similarity 36.7%; Prod. No. 5,2e-13;
 Matches 51; Conservative 17; Mismatches 55; Indels 16; Gaps 6,
 QY 2 ISGNSPPILNCRISYYS---TPIAVTVIVYS-----SGTRLIGKSLLCITKDKV 52
 DB 1517 ISCEPPTISNG--DFYSNNPTSPHNTVVYQHTGPDGEQLFELVGERSTYTSKDQ 1574
 QY 53 DETWIKAPKTEYENKYSSCEPIVPGYKLRGSLP-YRHCDSVTPACKTFNFSMGNKSV 111
 DB 1575 VGVWSSPPPTISTNK---CTAPEVFNATIRVFNPSFSLTETIRFQGFVWVGSHTV 1531
 QY 112 WQANNMNGPTRLPTCVSV 130
 DB 1642 QGQINGRWGP KLPHCSRV 1649
 RESULT 7
 AB010287
 ID AB000287 standard; Protein; 2039 AA.
 XX AC AB000287;
 XX PR 14 FEB 2002 (first entry)
 XX PR Novel human diagnostic protein #278.
 XX PR

KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
 KW food supplement; medical imaging; diagnostic; genetic disorder.
 XX OS Homo sapiens.
 XX PR WO200175067-A2.
 XX PR 11-OCT-2001.
 XX PR 30-MAR-2001; 2001WO-0508631.
 XX PR 31-MAR 2000; 2000US 0540217.
 XX PR 23-AUG-2000; 2000US-0649167.
 XX PR (HYSE-) HYSEQ INC.
 XX PR Drmanac RT, Liu C, Tanq YT.
 XX PR WI: 2001-639362/73.
 XX PR N PSDB; AAS64474
 XX PR New isolated polynucleotide and encoded polypeptides, useful in
 XX PR diagnostics, forensics, gene mapping, identification of mutations
 XX PR responsible for genetic disorders of other traits and to assess
 XX PR biodiversity
 XX PR Claim 20; SEQ ID No 30646; 10pp; English.
 XX CC The invention relates to isolated polynucleotide (I) and
 XX CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
 XX CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
 XX CC and gene mapping, and in recombinant production of (II). The
 XX CC polynucleotides are also used in diagnostics as expressed sequence tags
 XX CC for identifying expressed genes. (I) is useful in gene therapy techniques
 XX CC to restore normal activity of (II) or to treat disease states involving
 XX CC (II). (II) is useful for generating antibodies against it, detection or
 XX CC quantitating a polypeptide in tissue, as molecular weight markers and as
 XX CC a food supplement. (II) and its binding partners are useful in medical
 XX CC imaging of sites expressing (II). (I) and (II) are useful for treating
 XX CC disorders involving aberrant protein expression or biological activity.
 XX CC The polypeptide and polynucleotide sequences have applications in
 XX CC diagnostics, forensics, gene mapping, identification of mutations
 XX CC responsible for genetic disorders or other traits to assess biodiversity
 XX CC and to produce other types of data and products dependent on DNA and
 XX CC amino acid sequences. AB00010-ARG3-47 represent novel human
 XX CC diagnostic amino acid sequences of the invention.
 XX CC Note: The sequence data for this patent did not appear in the printed
 XX CC specification, but was obtained in electronic format directly from who
 XX CC at http://www.int/pdb/published/pdb_sequences.
 XX SQ Sequence 2039 AA;
 Query Match 29.2%; Score 219; DB 20; Length 2039;
 Best Local Similarity 36.7%; Prod. No. 5,2e-13;
 Matches 51; Conservative 17; Mismatches 55; Indels 16; Gaps 6,
 QY 2 ISGNSPPILNCRISYYS---TPIAVTVIVYS-----SGTRLIGKSLLCITKDKV 52
 DB 1517 ISCEPPTISNG--DFYSNNPTSPHNTVVYQHTGPDGEQLFELVGERSTYTSKDQ 1574
 QY 53 DETWIKAPKTEYENKYSSCEPIVPGYKLRGSLP-YRHCDSVTPACKTFNFSMGNKSV 111
 DB 1575 VGVWSSPPPTISTNK---CTAPEVFNATIRVFNPSFSLTETIRFQGFVWVGSHTV 1531
 QY 112 WQANNMNGPTRLPTCVSV 130
 DB 1642 QGQINGRWGP KLPHCSRV 1649
 RESULT 3
 AB010287
 ID AB=11782 standard; peptido; 2044 AA
 XX PR

FI /note- "See note a in comments below."
 FI 202
 FT /note- "See note a in comments below."
 XX
 PN USN729208 N.
 XX
 PD 14 MAR 1989.
 XX
 PE 20 AUG 1988; 8805 0249208.
 XX
 PR 20 AUG 1988; 8805 0249208.
 XX
 PA (USSH) NAT INST OF HEALTH.
 XX
 PI Korea G.
 XX
 DP WPI: 1989 165451/22.
 DP N-PSDB: AAN90113.
 XX
 PT New protein with anti-complement activity
 FT encoded by Vaccinia virus 35K gene
 XX
 PS Disclosure: Figure 2A: 20pp; English.
 XX
 CC 34b binding protein which specifically blocks human complement cascades.
 CC It is the deduced sequence of a 35kDa protein encoded by sequence 52-840
 CC of the 35K gene of vaccinia virus strain WR. Note a - these sites
 CC indicate the start of 60 amino acid tandem repeating units which have a
 CC consensus sequence. The signal peptide sequence is not found in purified
 CC 35K protein recovered from the medium of cells infected with vaccinia
 CC virus strain WR. A suggested use is to treat diseases due to abnormally
 CC high complement activity.
 CC (Note: Revised entry submitted to correct the patent number format of
 CC US government owned NTIS applications to prevent clashes with ongoing US
 CC granted patent numbers. For further information please visit the Derwent
 CC web site at www.derwent.com/dep/updates/ntis_us.html.)
 XX
 SQ Sequence 263 AA;
 Query Match 28.6%; Score 215; DB 10; Length 263;
 Best Local Similarity 35.9%; Pred. No. 1,1e-13;
 Matches 46; Conservative 15; Mismatches 55; Indels 12; Gaps 5;
 QY 2 ISGSSPPPLNGPISYSTPIAVITVIVYS:SGIERLLIGKSLIITKDKVETWKPAP 61
 DB 146 VKQSPSPISNGRHNGYEFETWTSVYSCNSGYSLIGSGVLT---SGGEWSDP-P 199
 QY 62 KQEFNKYSSTPEPIVPGYKIRG-STPYRGRD-VTFACKINFSMNGKNSVWQANNMWG 120
 DB 200 TQIVV---KQPIPTISNGYLSGGPKRYSYNDVDFCKRYKLSGSSSTCSGNTWK 255
 QY 121 PTRLPTCV 128
 DB 256 P-ELPKCV 262
 QY 121 PTRLPTCV 128
 DB 256 P-ELPKCV 262
 RESULT 14
 AAY29859
 ID AAY29859 standard; protein: 263 AA.
 XX
 AC AAY29859;
 XX
 DT 16 NOV 1999 (first entry)
 XX
 DE Vaccinia complement control protein sequence.
 XX
 KW Vaccinia virus; smallpox inhibitor of complement enzyme; SPIKE;
 KW fusion protein; hyperacute rejection; xenograft; inflammation;
 KW post ischemic reperfusion injury; malignancy; autoimmune disease;
 KW immune system disorder; neurodegeneration; infection; gene therapy;
 KW blood additive; extracorporeal circulation system.
 XX
 OS Vaccinia virus.

OS Synthetic.
 XX
 PN W09544625-A1.
 XX
 PD 10-SEP-1999.
 XX
 PE 02-MAR-1999; 99WO-US04635.
 XX
 PR 03-MAR-1998; 98TS-0076821.
 XX
 PA (UYJO) UNIV CHINS BOPKINS.
 PA (UYDI-) UNIV PITTSBURGH.
 XX
 PI Rosengard AM, Abecorn JM;
 XX
 DP WPI: 1999-550981/46.
 DP N-PSDB: AAZ21091.
 XX
 PT New smallpox inhibitor of complement enzyme protein, used to treat
 FT complement-mediated disease, particularly γ hyperacute rejection
 XX
 PS Claim 15; F.g 1; 88pp; English.
 XX
 CC The present invention describes the Vaccinia virus smallpox inhibitor of
 CC complement enzymes (SPIKE) protein. SPIKE is an inhibitor of complement
 CC activation, and so can be used to treat or prevent complement-mediated
 CC disorders, especially hyperacute rejection, inflammation or post
 CC ischemic reperfusion injury, malignancies, autoimmune diseases,
 CC immune system disorders, neurodegeneration and infections. Hyperacute
 CC rejection may also be prevented by treating the graft with SPIKE before
 CC transplanting it or by using a xenograft that has been transfected to
 CC express SPIKE from a gene therapy vector. SPIKE is also useful as
 CC additive to blood, e.g. in an extra-corporeal circulation system (coated
 CC on tubing) or in storage, also for studying complement activation.
 CC Transgenic animals that express SPIKE are used as sources of xenografts.
 CC The present sequence represents a vaccinia complement control protein
 CC (VCP) encoded by the specifically cloned mutated VCP nucleotide
 CC sequence, having a silent T to A transversion at nucleotide position
 CC number 267.
 XX
 SQ Sequence 263 AA;
 Query Match 28.6%; Score 215; DB 20; Length 263;
 Best Local Similarity 35.9%; Pred. No. 1,1e-13;
 Matches 46; Conservative 15; Mismatches 55; Indels 12; Gaps 5;
 QY 2 ISGSSPPPLNGPISYSTPIAVITVIVYS:SGIERLLIGKSLIITKDKVETWKPAP 61
 DB 146 VKQSPSPISNGRHNGYEFETWTSVYSCNSGYSLIGSGVLT---SGGEWSDP-P 199
 QY 62 KQEFNKYSSTPEPIVPGYKIRG-STPYRGRD-VTFACKINFSMNGKNSVWQANNMWG 120
 DB 200 TQIVV---KQPIPTISNGYLSGGPKRYSYNDVDFCKRYKLSGSSSTCSGNTWK 255
 QY 121 PTRLPTCV 128
 DB 256 P-ELPKCV 262
 RESULT 15
 AAB13014
 ID AAB13014 standard; protein: 263 AA.
 XX
 AC AAB13014;
 XX
 DT 11-DEC-2000 (first entry)
 XX
 DE Complement inhibitory protein, VCP amino acid sequence.
 XX
 KW Alzheimer's disease; Vaccinia virus VCP; complement pathway inhibitor;
 KW treatment; diagnosis; amyloid plaque
 XX
 OS Vaccinia virus.

XX Sequence 2049 AA;
 Query Match 28.2%; Score 212; DB 14; Length 2039;
 Best Local Similarity 36.0%; Pred. No. 2.6e 12;
 Matches 50; Conservative 19; Mismatches 54; Indels 16; Gaps 7;
 QY 2 ISGTSPPDILNGRIISVYS--TPIAVGIVIRYS-----SGPRILGKSKSLGCTPKDKV 52
 DB 1517 ISC EPPPTISNR DFYSNNKTSFUNGIVTYQCHITGPDGEQLQVRSISYCTSKDQ 1574
 QY 53 IGTWKPAPKTEYFNKYSSTPEPIVPGAYFIRIPITP YZHSVYTA/KTNFSNNKSV 111
 DB 1675 VGVWSSSPDRCISTNK--CTAENVQNAIRVGNRSFSETETIRPRCPGCFVWVSHIV 1631
 QY 112 WQANNMNGPTRLPTCVSV 130
 DB 1632 QQTNRWGF KLPICSRV 1649
 RESULT 17
 AAY55755
 ID AAY55755 standard; Protein: 450 AA.
 XX AAY55755;
 XX 22 FEB 2000 (first entry)
 DE Human CR1 protein LHR D SCR fragment.
 XX C3b/C4b receptor; CR1 protein; cell surface protein; erythrocyte; human;
 KW complement regulatory activity; complement pathway enzyme; tissue damage;
 KW reperfusion injury; Arthus reaction; myocardial infarct; inflammation;
 KW heart condition; autoimmune disorder; long homologous repeat; LHR; SCR;
 KW short consensus repeat.
 XX Homo sapiens.
 XX US'981481 A.
 XX 09 Nov 1999.
 XX 06 JUN 1995; 9505-0470652.
 XX 03 APR 1989; 8905 0332865.
 XX 06 DEC 1974; 7405-0350248.
 XX 24 FEB 1993; 9405-0026144.
 XX 01 APR 1988; 8805-0176542.
 XX (UYJO) UNIV JOHNS HOPKINS.
 XX (BHM) BRIGHAM & WOMENS HOSPITAL.
 XX (AVAN) AVANT IMMUNOTHERAPEUTICS INC.
 XX Concino MF, Wong WW, Makrides SC, Klickstein LB, Fearon DF, Ip SH;
 P Marsh HC, Carlson GR;
 XX WPI: 1999-63351/54.
 XX A human C3b/C4b receptor (CR1) protein having antiinflammatory and
 PT cardiant activity.
 XX Disclosure; Fig 10; 87pp; English.
 CC The invention relates to a human C3b/C4b receptor (CR1) protein. The CR1
 protein or fragment is expressed as a cell surface protein on the surface
 of a non human cell and exhibits a complement regulatory activity of full
 length human CR1 as expressed on erythrocytes. The CR1 function in vivo
 may be mediated through the inhibition of complement pathway enzymes. The
 soluble CR1 protein exhibits a complement regulatory activity, and thus
 may be used to prevent reperfusion injury, inhibit Arthus reaction, and
 neutrophil mediated tissue damage, and reduce myocardial infarct size,
 and inflammation. The CR1 protein and its fragments can also be used in
 the treatment of conditions which involve unwanted complement activity.

CC e.g., shock lung, tissue damage due to burn, or ischemic heart conditions,
 CC and autoimmune disorders. CR1 proteins, analogues, derivatives, and anti-
 CC -CR1 antibodies are used in assays, and diagnostics. The present sequence
 CC represents the short consensus repeat (SCR) fragments of human CR1
 CC protein long homologous repeat (LHR) sequence.
 XX Sequence 450 AA;
 Query Match 28.1%; Score 211; DB 40; Length 450;
 Best Local Similarity 36.0%; Pred. No. 5.2e 13;
 Matches 50; Conservative 17; Mismatches 56; Indels 16; Gaps 6;
 QY 2 ISGTSPPDILNGRIISVYS--TPIAVGIVIRYS-----SGPRILGKSKSLGCTPKDKV 52
 DB 124 ISCEPPTISNG--DFYSNNKTSFUNGIVTYQ YHGEAEQLLELWEEKSVCTSKDQ 191
 QY 53 IGTWKPAPKTEYFNKYSSTPEPIVPGAYFIRIPITP YZHSVYTA/KTNFSNNKSV 111
 DB 182 VGVWSSSPDRCISTNK--CTAENVQNAIRVGNRSFSETETIRPRCPGCFVWVSHIV 248
 QY 112 WQANNMNGPTRLPTCVSV 130
 DB 239 QQTNRWGF KLPICSRV 256
 RESULT 18
 AAY29858
 ID AAY29858 standard; Protein: 263 AA.
 XX AAY29858;
 XX 15 NOV 1999 (first entry)
 DE Vaccinia virus SPICE protein.
 XX Vaccinia virus.
 KW Vaccinia virus, smallpox inhibitor of complement enzyme; SPICE;
 KW fusion protein; hyperacute rejection; xenotraft; inflammation;
 KW post-ischemic reperfusion injury; ne injury; autoimmune disease;
 KW immune system disorder; neurodegeneration; infection; gene therapy;
 KW blood additive; extracorporeal circulation system.
 XX Vaccinia virus.
 XX Key location/Qualific's
 FT Misc-difference 13
 XX /note- "encoded by CNA"
 XX W09944625 A1.
 XX 10-SEP-1999.
 XX 02-MAR-1999; 99WO-US04635.
 XX 03-MAR-1998; 98US-0076821.
 XX (UYJO) UNIV JOHNS HOPKINS.
 XX (UYPI) UNIV PITTSBURGH.
 XX Rosequard AM, Ahearn JM;
 XX WPI: 1999-550981/46.
 XX N-PSDB; AAZ21690.
 XX New smallpox inhibitor of complement enzyme protein, used to treat
 FT complement-mediated disease, particularly hyperacute rejection
 XX Claim 1; Fig 2; 88pp; English.
 CC The present sequence represents Vaccinia virus smallpox inhibitor of
 CC complement enzymes (SPICE) protein. SPICE is an inhibitor of complemen
 CC activation, and so can be used to treat or prevent complement mediated
 CC disorders, especially hyperacute rejection, inflammation or fest,
 CC ischaemic reperfusion injury, malot, notes, autoimmune diseases.

Seq	Sequence	133 AA;	Matches	51;	Conservative	14;	Mismatches	57;	Indels	12;	Gaps	6;
Query Match	26.4%; Score 198; DB 15; Length 133;											
Best Local Similarity	48.1%; Prod. No. 2,44-12;											
Matches	51; Conservative	14; Mismatches	57; Indels	12; Gaps	6;							
QY	2 ISGSPPTLNGR EYYSIPIAVGVIRYSC--SG--TFELIGKSLILTIKKVVG 54		QY	2 ISGSPPTLNGR--ISYSTPIAVGVIRYSC--SG--TFELIGKSLILTIKKVVG 54								
DB	3 IPGLPTTNGDLSINPENNIGSVVYTPNPGSGKVFELVGP-IPYTSNGLQV 62		DB	1.4 IPGLPTTNGDLSINPENNIGSVVYTPNPGSGKVFELVGP-IPYTSNGLQV 62								
QY	5 TWDFAPKGFYKNSCPPIVGVYKIPGS--IPYDIPSVTEAFINPMGNKNSVW 113		QY	5 TWDFAPKGFYKNSCPPIVGVYKIPGS--IPYDIPSVTEAFINPMGNKNSVW 113								
DB	63 IWSGAPAGCIIPNK CTDFPVENGLIVSDNKSLSFNEVVERCQPGFVMKGPDRVRC 119		DB	63 IWSGAPAGCIIPNK CTDFPVENGLIVSDNKSLSFNEVVERCQPGFVMKGPDRVRC 119								
QY	114 GANNKSGTFLPTC 127		QY	114 GANNKSGTFLPTC 127								
DB	120 QALNKWEP ELPSC 132		DB	241 QALNKWEP ELPSC 253								
FASTA 28			RESULT 29									
AA047154			AA055754									
AA047154 standard; protein: 254 AA.			AA055754 standard; protein: 450 AA									
AA047154:			AA055754:									
14 JUN 1994 (first entry)			22 FEB 2000 (first entry)									
Sequence of soluble complement receptor type 1 (CRI) derivative			Human CRI protein LHR-C SCR fragment.									
which comprises the N terminal fragment M01-S253 of CRI.												
Complement receptor type 1 (CRI); short consensus repeat;												
long homologous repeat; domain: inflammation; therapy.												
Synthetic.												
W0400571 A.												
06 JAN 1994.												
16 JUN 1993; 9360 GH01282.												
24 JUN 1992; 9208 0013476.												
01 MAR 1993; 9308-0004057.												
(SMK) SMITHKLINE BEECHAM PLC.												
Dodd L. Freeman AM, Mossakowski DEL, Smith RM;												
WPI: 1994 026208/03.												
New soluble complement receptor type 1 deriva... used for												
treatment disease of disorder associated with inflammation or												
inappropriate complement activation												
Claim 15; Page 51-52; 65pp; English.												
CRI is composed of 40 short consensus repeats (SCRs) that each												
contain around 60-70 AAs. It is further arranged as 4 long												
homologous repeats (LHRs) of 7 SCRs each. Following a leader												
sequence, the CRI molecule consists of the N terminal LHR A, the												
next two repeats, LHR B and LHR C, and the most C-terminal LHR-D,												
followed by 2 additional SCRs, a 25 residue putative transmembrane												
region and a 43 residue cytoplasmic tail. Based on the mature CRI												
molecule, having a predicted N-terminal Glu, designated residue 1,												
the first four SCR domains of LHR-A are 2-58, 63-120, 125-191 and												
197-252. Of mature CRI, soluble fragments of CRI which corresp. to												
part of CRI possess functional complement inhibitory, including												
anti haemolytic activity.												
Sequence 254 AA:												
Query Match	26.4%; Score 198; DB 15; Length 254;											
Best Local Similarity	48.1%; Prod. No. 5,40-12;											
Matches	51; Conservative	14; Mismatches	57; Indels	12; Gaps	6;							

The invention relates to a human C3b/C4b receptor (CRI) protein. The CRI protein of fragment is expressed as a cell surface protein on the surface of a non-human cell and exhibits a complement regulatory activity of full-length human CRI as expressed on erythrocytes. The CRI function in vivo may be mediated through the inhibition of complement pathway enzymes. The soluble CRI protein exhibits a complement regulatory activity, and thus may be used to prevent reperfusion injury, inhibit Arthus reaction, and neutrophil-mediated tissue damage, and reduce myocardial infarct size, and inflammation. The CRI protein and its fragments can also be used in the treatment of conditions which involve unwanted complement activity, e.g. shock lung, tissue damage due to burn, or ischemic heart conditions, and autoimmune disorders. CRI protein, analogues, derivatives, and anti-CRI antibodies are used in assays and diagnostics. The present sequence represents the short consensus repeat (SCR) fragments of human CRI protein long homologous repeat (LHR) sequence.


```

28 APR 1992: 92EP 0304826.
03 MAY-1991: 91US 0695514
(UNIV ) UNIV WASHINGTON.
Atkinson JP, Bourcade D, Krych M;
WPI; 1992 375009/46.
Complement activity regulator protein analogues - useful for
treating autoimmune diseases, to suppress transplant rejection,
for diagnosis etc.
Claim 11: Fig 2 and R11810; 23pp; English.
The cDNA clone designated CRI-4 was described in J.Exp.Med.(1988)
168:1255-1270. It encodes the first 8 and a half amino terminal
SCRs of CRI. The invention concerns analogues of "regulator of
complement activation" proteins or truncated, hybrid or recombinant
forms of them. CRI-4 is a preferred truncated form and a number of
specified substitution variants of it are claimed. Positions 35 and
37 of SCR 1 and the corresponding positions in SCR 8 have been
identified as important in C4b binding. The specification does not
contain the CRI 4 sequence; the sequence given here was constructed
from the full length CRI amino acid sequence having GENESEQ
accession number AAR11810 and descriptions in the disclosure.
Sequence 543 AA;

Query Match 26.1%; Score 196; DB 13; Length 543;
Best Local Similarity 38.0%; Pred. No. 2.2e-11;
Matches 52; Conservative 14; Mismatches 59; Indels 12; Gaps 6;

QY 2 ISGSHPPPIINGR-LSYVSTPIAGVIVRSG- SG- TFRIGKSLIITKQKVDG 54
II III III III III III III III III III III III III III III
DB 123 IPGCLPPTITNGDFISTNBNFNHYGVSVTWYN-SSQAPWVFIVGELSYTSNGLWG 182
II III III III III III III III III III III III III III III
QY 55 TWKPAKPYEYKNYSQPEPIVPGYKIPQS- PYBPHSVTEA-KTNISNNKNSVVC 113
II III III III III III III III III III III III III III III
DB 183 IWSGAPAQCTIPNK---CTPPNVENGILVSDNKSLEINVEVFRCPVFMKGPRKVKC 239
II III III III III III III III III III III III III III III
QY 114 QANNMWGPTRLPTCVSV 130
II III III III III III III III III III III III III III III
DB 240 QALNKWEP-ELFSCSRV 255
II III III III III III III III III III III III III III III

RESULT 46
AAR28546
ID AAR28546 standard; peptide; 543 AA.
AC AAR28546;
QY 19 MAR 1993 (first entry)
DE CRI 4 (44T, 47D, 49L) analogue.
KW short consensus repeat; regulator of complement activation;
KW C4b binding; C4b binding; human complement type 1 receptor.
OS Homo sapiens.
PH Key location/Qualifiers
FT Region 1..60 /label= SCR 1
FT Region 61..122 /label= SCR 2
FT Region 451..510 /label= SCR 8
FT Region 511..543 /label= SCR 9
FT /note= "TRUNCATED"
FT Misc difference 44
```

```

FT Misc-difference 47 /note= "Ile substituted by Thr (SCR 8)"
FT /note= "Lys substituted by Asp (SCR 8)"
FT Misc-difference 49 /note= "Ser substituted by Leu (SCR 8)"
XX EP5:2733-A.
XX 11-NOV-1992.
XX 28-APR-1992: 92EP-0304826.
XX 03-MAY-1991: 91US-0695514.
XX (UNIV ) UNIV WASHINGTON.
XX Atkinson JP, Bourcade D, Krych M;
XX WPI; 1992-375009/46.
XX Complement activity regulator protein analogues - useful for
XX treating autoimmune diseases, to suppress transplant rejection,
XX for diagnosis etc.
XX Claim 11: Fig 2 and R11810; 23pp; English.
XX The cDNA clone designated CRI-4 was described in J.Exp.Med.(1988)
XX 168:1255-1270. It encodes the first 8 and a half amino terminal
XX SCRs of CRI. The invention concerns analogues of "regulator of
XX complement activation" proteins or truncated, hybrid or recombinant
XX forms of them. CRI-4 is a preferred truncated form and a number of
XX specified substitution variants of it are claimed in which certain
XX positions in SCR 1 which have been identified as important for the
XX degree of C3b and C4b-binding are substituted by amino acids from
XX the corresponding positions in SCR 8. The specification does not
XX contain the CRI-4 sequence; the sequence given here was constructed
XX from the full-length CRI amino acid sequence having GENESEQ
XX accession number AAR11810 and descriptions in the disclosure
XX Sequence 543 AA;

Query Match 26.1%; Score 196; DB 13; Length 543;
Best Local Similarity 38.0%; Pred. No. 2.2e-11;
Matches 52; Conservative 14; Mismatches 59; Indels 12; Gaps 6;

QY 2 ISGSHPPPIINGR-LSYVSTPIAGVIVRSG- SG- TFRIGKSLIITKQKVDG 54
II III III III III III III III III III III III III III III
DB 123 IPGCLPPTITNGDFISTNBNFNHYGVSVTWYN-SSQAPWVFIVGELSYTSNGLWG 182
II III III III III III III III III III III III III III III
QY 55 TWKPAKPYEYKNYSQPEPIVPGYKIPQS- PYBPHSVTEA-KTNISNNKNSVVC 113
II III III III III III III III III III III III III III III
DB 183 IWSGAPAQCTIPNK---CTPPNVENGILVSDNKSLEINVEVFRCPVFMKGPRKVKC 239
II III III III III III III III III III III III III III III
QY 114 QANNMWGPTRLPTCVSV 130
II III III III III III III III III III III III III III III
DB 240 QALNKWEP-ELFSCSRV 255
II III III III III III III III III III III III III III III

RESULT 37
AAR28547
ID AAR28547 standard; peptide; 543 AA
AC AAR28547;
XX 19-MAR-1993 (first entry)
XX CRI-4 (52S, 53S, 54P) analogue.
XX short consensus repeat; regulator of complement activation;
XX C3b binding; C4b binding; human complement type 1 receptor.
XX Homo sapiens
XX
```


GenCore version 5.1.1.3
Copyright (c) 1993-2002 CompuGen Ltd.

OM protein protein search, using sw model

Run on: November 6, 2002, 16:04:59 : Search time 14.4461 seconds
(without alignments)
891.311 Million cell updates/sec

Title: US 09 834 009 4

Perfect score: 751

Sequence: 1 G1SGVSPPTTIN:RISYST ANNMWGPTEPTCVSVFPLE 134

Scoring table: BLASTM6.2

Gapop 10.0, Gapext 0.5

Searched: 283148 seqs, 968044 residues

Total number of hits satisfying chosen parameters: 283148

Minimum DB seq length: 0

Maximum DB seq length: 20000000

Post processing: Minimum Match 0%

Maximum Match 100%

Listing first 100 summaries

Database:

PIR 71:*

1: pir1:*

2: pir2:*

3: pir3:*

4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	751	100.0	1091	1 P10009	complement C3d/Eps
2	466.5	62.1	363	2 B45900	complement C3d/Eps
3	466.5	62.1	676	2 A45900	complement C3d/Eps
4	466.5	62.1	1025	1 A45926	complement C3d/Eps
5	220	29.3	2489	2 I38012	complement C3b/C4b
6	218	29.0	2014	2 I36936	complement receptor
7	215	28.6	263	1 WW2231	apolipoprotein H h
8	210	28.0	263	1 C36848	complement control
9	210	29.0	263	2 I28456	hypothetical prote
10	208	27.7	263	2 B22352	B18L protein var
11	203.5	27.1	579	2 A56740	sperm egg recogniti
12	202.5	27.0	597	1 S53711	CAMP alpha chain p
13	201	26.8	482	2 A44924	complement C3b/C4b
14	199	26.5	597	1 NHH054	C4b-binding protei
15	197.5	26.3	610	1 I46001	C4b-binding protei
16	184	24.5	497	2 J62054	complement regulat
17	182	24.2	558	2 S57954	C4BP protein alpha
18	180.5	24.0	440	2 A45519	complement recepto
19	180.5	24.0	1231	1 NHH018	complement factor
20	178.5	23.8	469	1 NHH054	C4b-binding protei
21	178.5	23.8	1234	1 NHH054	complement factor
22	165.5	22.0	808	2 D45069	complement factor
23	165	22.0	661	1 KRU013	coagulation factor
24	164	21.8	2043	2 T18524	scavenger receptor
25	162	21.6	343	2 G35070	apolipoprotein H:
26	160	21.3	362	2 J65194	membrane cofactor
27	160	21.3	369	2 J65148	membrane cofactor
28	159	21.2	377	2 I54479	membrane cofactor
29	159	21.2	484	2 S01896	membrane cofactor

30	158.5	21.1	340	2	I56274	decay-accelerating
31	158.5	21.1	381	1	B26359	decay-accelerating
32	158.5	21.1	440	2	A26359	decay-accelerating
33	155	20.6	349	2	G02913	sperm C14b - human
34	155	20.6	369	2	I57998	membrane cofactor
35	150	20.0	302	1	WMBE1E	secretory compleme
36	150	20.0	360	1	WMBE2E	membrane bound com
37	149	19.8	668	2	A46013	coagulation factor
38	148	19.7	768	2	A42755	p-selectin precurs
39	147	19.6	330	2	I56100	complement factor
40	146.5	19.5	768	2	I53821	p-selectin - tat
41	142	18.9	1019	2	A48748	coagulation factor
42	141	18.8	252	2	A44877	C4b-binding protei
43	140	18.6	152	2	A53274	complement factor
44	140	18.6	258	2	S57960	C4BP protein beta
45	140	18.6	752	1	C28U	complement C2 prec
46	137	18.2	345	1	NHMS	apolipoprotein H p
47	135	18.0	360	2	T42921	complement control
48	134.5	17.9	270	2	I37278	complement factor
49	134	17.8	297	1	NHRT	apolipoprotein H p
50	134	17.8	763	2	I50807	complement factor
51	133	17.7	764	1	NHHU	complement factor
52	132	17.6	345	1	JN0465	apolipoprotein H p
53	131	17.4	202	1	A44247	C4b binding protei
54	131	17.4	669	2	S65551	factor II - bovine
55	129.5	17.2	702	2	T16842	hypothetical prote
56	128	17.0	560	2	I16843	hypothetical prote
57	127	16.9	345	1	NHHU	apolipoprotein H p
58	127	16.9	345	1	NHHU	apolipoprotein H p
59	127	16.9	646	2	JN0473	p-selectin precurs
60	125	16.6	868	2	T20249	hypothetical prote
61	125	16.6	1827	2	T34288	hypothetical prote
62	124.5	16.6	977	2	I52657	structure-related pr
63	123.5	16.4	612	2	B42755	p-selectin precurs
64	123	16.4	934	1	A44372	complement C6 prec
65	122	16.2	761	1	NHMS	complement factor
66	120.5	16.0	551	2	I46759	cadherin-like lecto
67	120.5	16.0	760	1	C2MS	classical compleme
68	120	16.0	330	2	I55975	X/Y protein - mous
69	120	16.0	610	2	A50316	p-selectin precurs
70	118	15.7	331	2	A45232	complement factor
71	118	15.7	482	2	J65092	E-selectin - pig
72	116.5	15.5	1797	2	T21899	hypothetical prote
73	116.5	15.5	1805	2	T21898	hypothetical prote
74	116	15.4	303	2	A35068	apolipoprotein H:
75	114	15.2	452	2	A35068	complement factor
76	114	15.2	830	2	A30359	p-selectin precurs
77	113	15.0	570	2	T46261	hypothetical prote
78	111.5	14.8	449	1	NHH054	complement factor
79	111.5	14.8	1506	2	I30886	integratory muc
80	111	14.8	747	2	I51579	complement factor
81	110.5	14.7	843	1	A27340	complement C7 prec
82	110	14.6	485	2	I36772	E-selectin - bovin
83	106.5	14.2	1053	2	S46199	probable compleme
84	102.5	13.6	317	2	A84933	ps/ur protein - va
85	101.5	13.5	686	1	A59271	Ca reactive factor
86	98.5	13.3	317	2	I47412	trypsin-activat
87	97.5	13.0	317	2	T28605	hypothetical prote
88	94.5	12.6	317	2	J61799	p6k protein precu
89	93	12.4	347	1	NHRT	apolipoprotein precu
90	91.5	12.3	317	2	T72172	HR protein - vari
91	91.5	12.2	317	2	G36855	HR protein - vari
92	90.5	12.1	317	2	I42526	p6k protein - vari
93	86	11.5	699	1	I54763	Ca-reactive factor
94	85.5	11.4	926	1	OPPGIT	radical peroxidase
95	85	11.3	198	2	I46002	C4BP beta chain
96	85	11.3	347	1	HPMS	apolipoprotein precu
97	85	11.3	633	2	T24898	hypothetical prote
98	84	11.2	370	2	S22124	p-selectin precurs
99	83.5	11.1	406	1	NHHU2	haptoglobin precu
100	83	11.1	933	1	OPPHIT	radical peroxidase

A:Experimental source: clone 411
 J. Exp. Med. 191, 151-159, 1995
 A:Title: Murine complement regulatory protein Crfy/p65 uses the specific mechanisms of bc
 A:Reference number: 148306; MUID:95105691
 A:Accession: 148306
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 21 467 -RES-
 A:Cross references: EMBL:U17128; NID:9595980; PIDN:AAA78271 1; PID:9595982
 C:Notes:
 A:Introns: 80/7; 113/2; 142/1; 275/1; 334/2
 C:Superfamily: complement factor H repeat homology
 F:22-78/Domain: complement factor H repeat homology <FH01>
 F:83-140/Domain: complement factor H repeat homology <FH02>
 F:145-211/Domain: complement factor H repeat homology <FH03>
 F:217-272/Domain: complement factor H repeat homology <FH04>
 F:276-331/Domain: complement factor H repeat homology <FH05>
 F:336-394/Domain: complement factor H repeat homology <FH06>
 F:399-458/Domain: complement factor H repeat homology <FH07>
 F:467-524/Domain: complement factor H repeat homology <FH08>
 F:531-587/Domain: complement factor H repeat homology <FH09>
 F:592-648/Domain: complement factor H repeat homology <FH10>
 Query Match 62.18; Score 466.5; DB 2; Length 676;
 Best Local Similarity 61.28; Pred No. 6,2e-37;
 Matches 82; Conservative 18; Mismatches 33; Indels 1; Gaps 1;
 QY 2 ISCGSPPTLNGRISYSTPTIAGVITVTP*SGSGTFRLEIGKSLICITK*VWGTWKPAP 61
 DB 497 ISCDPPVVKARKPYSLIPVPGTVLR*VNS*SYRL*GEKAIICISENOVIATWIKAPP 456
 QY 62 KCFYENKYSSGPETIVAGYKIRGS IYRHRGDSVTFACKTINFSMGNKSWW*QANNMWG 120
 DB 457 ICESVNVKITSQSDPIVPGFNNKSKAP*ERHGDSTVTP*CKANF*TMKSGKIVW*QANNMWG 516
 QY 121 PTRLFTCVSVFPLE 134
 DB 517 PTALVCHSDPPL 540
 RESULT 4
 A:Immunol. 148, 1458-1467, 1990
 A:Title: Comparative structure and evolution of murine CR2. The homolog of the human C3d
 A:Reference number: A43526; MUID:90229735
 A:Accession: A43526
 A:Molecule type: mRNA
 A:Residues: 1-1025 <FIN>
 A:Cross references: GR:M5684; EMBL:104153; NID:9192487; PIDN:AAA37448.1; PID:4192588
 R:Molina, B.; Kinoshita, T.; Inoue, K.; Garel, F.; Holers, V.M.
 J. Immunol. 145, 2974-2983, 1990
 A:Title: A molecular and immunological characterization of mouse CR2. Evidence for a si
 A:Reference number: A43538; MUID:91010789
 A:Accession: A43538
 A:Molecule type: mRNA
 A:Residues: 12-307, 519, 521-1025 <MOL>
 A:Cross references: GR:M1142; NID:9192492; PIDN:AAA63295.1; PID:4192593
 R:Flindt, J.D.; Benedict, M.A.; Levy, D.N.; Strominger, J.L.
 Proc. Natl. Acad. Sci. U.S.A. 86, 242-246, 1989
 A:Title: Identification of murine complement
 A:Reference number: A42215; MUID:89098890
 A:Accession: A42215
 A:Molecule type: mRNA
 A:Residues: 343-431/991 1025 -F12-
 A:Cross references: GR:104154

R:Kurtz, C.B.; Paul, M.S.; Acqter, M.; Weis, J.J.; Weiss, J.H.
 J. Immunol. 143, 2058-2067, 1989
 A:Title: Murine complement receptor gene family. Identification and characterization
 A:Reference number: A45802; MUID:89381590
 A:Accession: A45802
 A:Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 76, 100-101, 292-963, 964-1025 -F0R>
 A:Cross references: GR:M29281; NID:9192485; PIDN:AAA7447.1; PID:4087151
 A:Note: the authors failed to translate CCG for residue 421 as Gly, and CCA for resid
 C:Superfamily: complement C3d/Epstein-Barr virus receptor; complement factor H repeat
 C:Keywords: alternative splicing; duplication; glycoprotein; receptor; transmembrane
 F:1-11/Domain: signal sequence #status predicted -SIG-
 F:12-973/Domain: complement C3d/Epstein-Barr virus receptor 2 (15-repeat form) data
 F:14-73/Domain: complement factor H repeat homology <FH01>
 F:82-134/Domain: complement factor H repeat homology <FH02>
 F:146-202/Domain: complement factor H repeat homology <FH03>
 F:207-263/Domain: complement factor H repeat homology <FH04>
 F:268-334/Domain: complement factor H repeat homology <FH05>
 F:343-398/Domain: complement factor H repeat homology <FH06>
 F:402-458/Domain: complement factor H repeat homology <FH07>
 F:463-514/Domain: complement factor H repeat homology <FH08>
 F:519-585/Domain: complement factor H repeat homology <FH09>
 F:594-649/Domain: complement factor H repeat homology <FH10>
 F:654-704/Domain: complement factor H repeat homology <FH12>
 F:709-769/Domain: complement factor H repeat homology <FH13>
 F:778-833/Domain: complement factor H repeat homology <FH14>
 F:841-897/Domain: complement factor H repeat homology <FH15>
 F:902-958/Domain: complement factor H repeat homology <FH16>
 F:958-989/Domain: transmembrane #status predicted -TM-
 F:990-1025/Domain: intracellular #status predicted -INT-
 Query Match 62.18; Score 466.5; DB 1; Length 1025;
 Best Local Similarity 61.28; Pred No. 9,7e-37;
 Matches 82; Conservative 18; Mismatches 33; Indels 1; Gaps 1;
 QY 2 ISCGSPPTLNGRISYSTPTIAGVITVTP*SGSGTFRLEIGKSLICITK*VWGTWKPAP 61
 DB 12 ISCDPPVVKARKPYSLIPVPGTVLR*VNS*SYRL*GEKAIICISENOVIATWIKAPP 71
 QY 62 KCFYENKYSSGPETIVAGYKIRGS IYRHRGDSVTFACKTINFSMGNKSWW*QANNMWG 120
 DB 72 ICESVNVKITSQSDPIVPGFNNKSKAP*ERHGDSTVTP*CKANF*TMKSGKIVW*QANNMWG 516
 QY 121 PTRLFTCVSVFPLE 134
 DB 132 PTALVCHSDPPL 145
 RESULT 5
 173012
 N:Alternate names: complement C3b/C4b receptor; complement receptor type 1 (CR1); su
 N:Contains: complement C3b/C4b receptor; complement receptor type 1 (CR1); su
 C:Species: Homo sapiens (man)
 C:Date: 92 Jul 1996 #sequence revision 24 Rev:1999 #text_change 21-Jul-2000
 A:Accession: 173012; 156203; A47602; S13291; S03843; A28507; A24748; B24748; C24748
 R:Vik, D.P.; Wong, W.W.
 J. Immunol. 157, 6214-6224, 1993
 A:Title: Structure of the gene for the Fc allele of complement receptor type 1 and seq
 A:Reference number: 156203; MUID:94065175
 A:Accession: 173012
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-683, 685-1133, 1471, 1473, 1474, 2489 -VIR1-
 A:Cross references: GR:117418; NID:930674; PIDN:AA60695.1; PID:4451304
 A:Accession: 156203
 A:Status: translated from GB/EMBL/DDBJ
 A:Molecule type: DNA
 A:Residues: 1-683, 685-894, 896-1039 1451-1471, 1473, 2489 -VIR2-
 A:Cross references: GR:117418; NID:930673; PIDN:AA60694.1; PID:4406600

A:Title: Identification of an alternative polyadenylation site in the human C4b/C4b receptor type 1.

A:Reference number: S03291; MUID:89010527

A:Accession: S03292

A:Status: nucleotide acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 7975-8022

A:Cross references: EMBL: X14360

A:Note: The nucleotide sequence was submitted to the EMBL Data Library, November 1988

C:Superfamily: C4b-binding protein, alpha chain, complement factor H repeat homology

F:143 99/Domain: complement factor H repeat homology <PH01>

F:164 161/Domain: complement factor H repeat homology <PH02>

F:166 242/Domain: complement factor H repeat homology <PH03>

F:248 293/Domain: complement factor H repeat homology <PH04>

F:297 353/Domain: complement factor H repeat homology <PH05>

F:358 416/Domain: complement factor H repeat homology <PH06>

F:421 478/Domain: complement factor H repeat homology <PH07>

Query Match 26.89, Score 291, DB 2, Length 482.

Best Local Similarity 46.99, Pred. No. 1,3e-11;

Matches 50; Conservative 15; Mismatches 60; Indels 12; Gaps 5;

QY 2 TCGSPPHLEKRSVYKSTP-LAWGVIVFVSQS- ----GPERLGGKSLGCTKKQKVG 54

DB 164 HICGLPPTIANDFTSISREYTHGVSVVYTHCNLGRCKKVFELVGEPSIVCTSKRDQVG 223

QY 65 TWKDAIPKQYENKYSKDEPVPQGYKIRGS-TPYARHGDSVTPACKTNF-SMNCNKSVMC 113

DB 224 WSGFAVQVQFNK- - - - -GTPNVNLGTVLSVNRSLPSNLVVEPRQVGHMKGUSHVKC 280

QY 114 QANNMWGTRLPFTVSV 130

DB 281 QALRKWED-ELFSGSRV 296

RESULT 14

SDH004

C4b-binding protein alpha chain precursor - human

A:Alternative names: C4BP; proline rich protein

C:Species: Homo sapiens (man)

C>Date: 13-Aug-1986 #sequence

C:Accession: A33568, S02472, A0026, A24042, A34146, S24092, A31785, I52244, A04210

R:Matsumuchi, T.; Okamura, S.; Aso, T.; Sata, T.; Niho, Y.

Biochem. Biophys. Res. Commun. 165, 138-144, 1989

A:Title: Molecular cloning of the cDNA coding for proline rich protein (PRP): identity of

A:Reference number: A33568; MUID:90073699

A:Accession: A33568

A:Molecule type: mRNA

A:Residues: 1-597 <MAL>

A:Cross references: GR:M1452; NID:q160501; PIDW:AAA36507 1; PID:q190502

A:Note: The authors translated the codon GGA for residue 492 as Glu

Refining, S.J.; Lewin, A.R.; Reid, K.R.M.

FEBS Lett. 242, 328-332, 1988

A:Title: Derivation of the sequence of the signal peptide in human C4b-binding protein

A:Reference number: S02472; MUID:88242821

A:Accession: S02472

A:Molecule type: mRNA

A:Residues: 17-81 <L12>

A:Cross references: EMBL: X07853

A:Note: Although the sequence determined extends to residue 9 above, these authors assigned

R:Chung, L.P.; Bentley, D.R.; Reid, K.R.M.

Biochem. J. 250, 133-141, 1985

A:Title: Molecular cloning and characterization of the cDNA coding for C4b-binding protein

A:Reference number: A30426; MUID:86025405

A:Accession: A30426

A:Molecule type: mRNA

A:Residues: 80-597 <CL2>

A:Cross references: GR:X02865; NID:q20564; PIDW:CAA26517 1; PID:q20565

A:Note: 92 Thr and 457 His were also found

Refining, S.J.; Reid, K.R.M.

FEBS Lett. 204, 77-81, 1986

A:Title: Studies on the structure of the human C4b-binding protein gene.

A:Reference number: A24182; MUID:8640119

A:Accession: A24182

A:Molecule type: DNA

A:Residues: 203-288 <LIN>

A:Cross references: EMBL:X04284; EMBL:X05296

R:Rodriguez de Cordoba, S.; Sanchez-Corral, F.; Rey-Campos, J.

J. Exp. Med. 173, 1073-1082, 1991

A:Title: Structure of the gene coding for the alpha polypeptide chain of the human C4b-binding protein

A:Reference number: A33024; MUID:91217616

A:Contents: annotation; exon-intron boundaries

R:Chung, L.P.; Bentley, D.R.; Reid, K.R.M.

Mol. Immunol. 22, 427-435, 1985

A:Title: Amino acid sequence studies of human C4b-binding protein: N terminal sequence

A:Accession: A33134; MUID:8529600

A:Reference number: A33134

A:Molecule type: protein

A:Residues: 49-87 <CH1>

A:Note: This paper reports amino terminal sequences of the intact protein and of a sub

F:163-164, M.; Kato, S.; Takeya, H.; Yam, Y.; Veer, C.; Harkov, T.M.; Iwanaga, S.; B

FEBS Lett. 317, 228-232, 1993

A:Title: The region Sor(333)-Arg(356) of the alpha chain of human C4b-binding protein

A:Reference number: S29492; MUID:9314616

A:Accession: S29492

A:Status: preliminary

A:Molecule type: protein

A:Residues: 381-404 <IES>

R:Suzuki, K.; Nishio, J.

J. Biol. Chem. 263, 17034-17039, 1988

A:Title: Binding site for vitamin K-dependent protein S on complement C4b-binding pro

A:Reference number: A31785; MUID:89034204

A:Accession: A31785

A:Molecule type: protein

A:Residues: 495-505, X', 507-510, X', 512-515, S02>

A:Note: this peptide appears to bind protein S

R:Dahlback, H.; Smith, C.A.; Muller-Eberhard, H.J.

Proc. Natl. Acad. Sci. U.S.A. 80, 3461-3465, 1983

A:Title: Visualization of human C4b-binding protein and its complexes with vitamin K-

A:Reference number: A33950; MUID:84221615

A:Contents: annotation; electron microscopy, three dimensional structure; ligand bind

R:Aso, T.; Okamura, S.; Matsuguchi, T.; Sakamoto, N.; Sata, T.; Niho, Y.

Biochem. Biophys. Res. Commun. 174, 222-227, 1991

A:Title: Genomic organization of the alpha chain of the human C4b-binding protein gen

A:Reference number: I52244; MUID:91113199

A:Accession: I52244

A:Status: translation not shown, translated from GR/EMBL/DBD

A:Molecule type: DNA

A:Residues: 1-597 <ASO>

A:Cross references: GR:M62486; NID:q150498; PIDW:AAA36506 1; PID:q190560

C:Comment: C4BP controls the classical pathway of complement activation. It binds as

the degradation of the C4bC2a complex (C3 convertase) by dissociating the complement

C:Comment: C4BP occurs in plasma in two forms, both of which bind complement fragment

isulfide bonds. The beta chain binds the vitamin K dependent plasma protein S. Amino

actors V and VIII.

C:Comment: The molecule has a central body affected by seven tentacles (alpha chains).

C:Genetics:

A:Gene: GDB:C4BPA

A:Cross references: GR:120568; OMIM:120830

A:Map position: 1q12.1q32

A:Features: 48/1, 137/1, 143/2, 172/1, 230/1, 247/1, 362/1, 425/2, 482/2, 493/4

C:Complex: octamer of seven alpha chains and one beta chain

C:Superfamily: C4b-binding protein alpha chain; complement factor H repeat homology

C:Keywords: acute phase; chylomicron; complement pathway; duplication; glycoprotein;

F:148/Domain: signal sequence #status predicted <SIG>

F:49-597/Domain: C4b binding protein alpha chain #status predicted <MAL>

F:50-108/Domain: complement factor H repeat homology <PH1>

F:113 170/Domain: complement factor H repeat homology <PH2>

F:175 234/Domain: complement factor H repeat homology <PH3>

F:235 291/Domain: complement factor H repeat homology <PH4>

F:293 360/Domain: complement factor H repeat homology <PH5>

F:361 422/Domain: complement factor H repeat homology <PH6>

F:381-404/Region: complement C4b binding #status predicted <IES>

F:426-483/Domain: complement factor H repeat homology <PH7>

F:484-539/Domain: complement factor H repeat homology <PH8>

A.Molecule type: mRNA
A.Residues: 1-369 <MUR>
A.Cross references: DDBJ:p64811; NID:q1777415; PID:q1777415
C.Comment: This protein is a complement regulator and measles virus receptor. It protect
C.Superfamily: herpesvirus complement control protein; complement factor H repeat homolo
F.1-34/Domain: signal sequence #status predicted <SIG>
F.35-94/Domain: membrane cofactor protein (MCP, CD46) receptor
F.99-157/Domain: complement factor H repeat homology <PH01>
F.162-224/Domain: complement factor H repeat homology <PH02>
F.228-283/Domain: complement factor H repeat homology <PH03>
F.329-351/Domain: complement factor H repeat homology <PH04>
F.347-349/Domain: transmembrane #status predicted <TM>
F.347-349/Domain: intracellular #status predicted <IM>

Query Match 21.2% Score 160 DB 2 Length 369
Best Local Similarity 29.1% Pred. No. 8 3e-08
Matches 49 Conservative 22 Mismatches 57 Indels 16 Gaps 6

QY 2 ISGSPPLNGRISYSSTPI AVGVIVYKTRG STPYRHCDSVTFAC KINFSMNKSWVWQ 114
DB 160 ILCTTPPKIKNGKHTESEVEVEYLDVAVYSCDPAPDPDPESLIGESMIVGN----NST 215

QY 56 WDKPAKKEEYFNKYSSTPI AVGVIVYKTRG STPYRHCDSVTFAC KINFSMNKSWVWQ 114
DB 216 WSHAAPEG----KVKKKEEYFNKYSSTPI AVGVIVYKTRG STPYRHCDSVTFAC KINFSMNKSWVWQ 271

QY 115 ANNMGPTRLPCTV 128
DB 272 SNSTWDP-VPKCKV 284

RESULT 28
154479
membrane cofactor protein precursor, splice form p64.1 human
N.Alternate names: lymphocyte surface glycoprotein CD46
C.Species: Homo sapiens (man)
C.Date: 07-Jun-1996 #sequence_revision 07-Jun-1996 #text_change 31-Mar-2000
C.Accession: 154479
A.Title: Alternatively spliced mRNAs encode several isoforms of CD46 (MCP), a regulator o
A.Reference number: 154479; MUID:91267562
A.Cross references: GDB:M88050; NID:q180136; PID:AAA628331; FID:q180137
A.Coverages: 154479
A.Molecule type: mRNA
A.Residues: 1-377 <RES>
A.Cross references: GDB:M88050; NID:q180136; PID:AAA628331; FID:q180137
A.Coverages: 154479
A.Molecule type: mRNA
A.Residues: 1-369 <MUR>
A.Cross references: DDBJ:p64811; NID:q1777415; PID:q1777415
C.Comment: This protein is a complement regulator and measles virus receptor. It protect
C.Superfamily: herpesvirus complement control protein; complement factor H repeat homolo
F.1-34/Domain: signal sequence #status predicted <SIG>
F.35-94/Domain: membrane cofactor protein (MCP, CD46) receptor
F.99-157/Domain: complement factor H repeat homology <PH01>
F.162-224/Domain: complement factor H repeat homology <PH02>
F.228-283/Domain: complement factor H repeat homology <PH03>
F.329-351/Domain: complement factor H repeat homology <PH04>
F.347-349/Domain: transmembrane #status predicted <TM>
F.347-349/Domain: intracellular #status predicted <IM>

Query Match 21.2% Score 159 DB 2 Length 377
Best Local Similarity 27.9% Pred. No. 11e-07
Matches 48 Conservative 25 Mismatches 57 Indels 16 Gaps 6

QY 2 ISGSPPLNGRISYSSTPI AVGVIVYKTRG STPYRHCDSVTFAC KINFSMNKSWVWQ 114
DB 160 ILCTTPPKIKNGKHTESEVEVEYLDVAVYSCDPAPDPDPESLIGESMIVGN----NST 215

QY 56 WDKPAKKEEYFNKYSSTPI AVGVIVYKTRG STPYRHCDSVTFAC KINFSMNKSWVWQ 114
DB 216 WSHAAPEG----KVKKKEEYFNKYSSTPI AVGVIVYKTRG STPYRHCDSVTFAC KINFSMNKSWVWQ 271

QY 115 ANNMGPTRLPCTVSV 130
DB 272 SNSTWDP-VPKCKV 286

RESULT 29
S01896
membrane cofactor protein precursor - human
N.Alternate names: lymphocyte surface glycoprotein CD46
C.Species: Homo sapiens (man)
C.Date: 30-Sep-1995 #sequence_revision 03-Sep-1998 #text_change 21-Jul-2000
C.Accession: S01896; A60765; 156188
R.Lebien, D.M., Liszewski, M.K., Post, J.W., Aruffo, A., Le Beau, M.M., Reberthaus, M.
J. Exp. Med. 164, 181-194, 1988
A.Title: Molecular cloning and chromosomal localization of human membrane cofactor pr
A.Reference number: S01896; MUID:88266080
A.Coverages: S01896
A.Molecule type: mRNA
A.Residues: 1-384 <LUR>
A.Cross references: FMO:Y00651; NID:q34501; PID:CAA06751; FID:q34505
A.Note: Part of this sequence, including the amino end of the mature protein, was con
R.Purcell, D.F.J., Deacon, N.J., Andrew, S.A., McKenzie, I.F.C.
Immunogenetics 31, 21-28, 1990
A.Title: Human non lineage antigen, CD46 (CD11b): purification and partial sequenc
A.Reference number: A60765; MUID:90129152
A.Coverages: A60765
A.Molecule type: protein
A.Residues: 36-39, 41-58 <PUR>
R.Cui, W., Boucade, D., Post, T., Greenhalgh, A.C., Atkinson, J.P., Kumar, V.
J. Immunol. 151, 4137-4146, 1993
A.Title: Characterization of the promoter region of the membrane cofactor protein (CD
A.Reference number: 156188; MUID:94014356
A.Coverages: 156188
A.Status: preliminary; translated from GEX/EMUL/10031
A.Molecule type: DNA
A.Residues: 1-34 <RES>
A.Cross references: GDB:S65879; NID:q425643; PID:AA0146681; FID:q425643
C.Coverages: 154479
A.Molecule type: mRNA
A.Residues: 1-369 <MUR>
A.Cross references: DDBJ:p64811; NID:q1777415; PID:q1777415
C.Comment: This protein is a complement regulator and measles virus receptor. It protect
C.Superfamily: herpesvirus complement control protein; complement factor H repeat homolo
F.1-34/Domain: signal sequence #status predicted <SIG>
F.35-94/Domain: membrane cofactor protein (MCP, CD46) receptor
F.99-157/Domain: complement factor H repeat homology <PH01>
F.162-224/Domain: complement factor H repeat homology <PH02>
F.228-283/Domain: complement factor H repeat homology <PH03>
F.329-351/Domain: complement factor H repeat homology <PH04>
F.347-349/Domain: transmembrane #status predicted <TM>

Query Match 21.2% Score 159 DB 2 Length 384
Best Local Similarity 27.9% Pred. No. 11e-07
Matches 38 Conservative 25 Mismatches 57 Indels 16 Gaps 6

QY 2 ISGSPPLNGRISYSSTPI AVGVIVYKTRG STPYRHCDSVTFAC KINFSMNKSWVWQ 114
DB 160 ILCTTPPKIKNGKHTESEVEVEYLDVAVYSCDPAPDPDPESLIGESMIVGN----NST 215

QY 56 WDKPAKKEEYFNKYSSTPI AVGVIVYKTRG STPYRHCDSVTFAC KINFSMNKSWVWQ 114
DB 216 WSHAAPEG----KVKKKEEYFNKYSSTPI AVGVIVYKTRG STPYRHCDSVTFAC KINFSMNKSWVWQ 271

QY 115 ANNMGPTRLPCTVSV 130
DB 272 SNSTWDP-VPKCKV 286

RESULT 30
156234
decay accelerating factor - orangutan (fragment)


```

1000 Local Similarity: 20.5% Prod. No. 1.5e 0%
Matches: 47/ Conservative 16/ Mismatches 61/ Indels 16/ Gaps 6/
QY 2 ENGGSEPTING RGVYSSTPLAVGLIVRYSSTPLRGLGKSLICITKDKVRLMD 57
DE 97 VIGPAVSTFENGLYIFRLASY FVGNVSFECDGGTLRGSVVRQ 148
QY 56 KIAPKCFEYKYSSTPEPTVDSYKIRGSTPYRIRDSVTEACRINFNNIKRSVRCVANN 117
ID 130 GELAVTD NGAGHCHNDVLSLQ AVRIQEREGHGRDKVRYRCSSNLVLGSSRETCUNG 195
QY 119 MRGPTDLPTN L 1
ID 11 1 1 1 1
ID 16 VWSGTF FIC 204

```

Search completed: November 6, 2002, 16:07:30
 Job time: 19.4164 sec

GenInfo version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd

OM protein protein search, using sw model

Run on: November 6, 2002, 16:04:58 : Search time 8.4684 Seconds
(without alignments)
612.680 Million cell updates/sec

Title: US 09-834-309-4

Perfect score: 761

Sequence: 1 G1SGSPPTLNGPISYST

ANMMWPTWCPVSVFPLE 134

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 105224 seqs, 4819550 residues

Total number of hits satisfying eblast parameters: 105224

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post processing: Minimum Match 0%

Maximum Match 100%

Listing first 100 summaries

Database : SwissProt_40.*

Prod. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query %	Length	DB ID	Description
1	751	100.0	1033	1	P20023 homo sapien
2	466.5	62.1	1025	1	CR2_MOUSE
3	219	29.2	2039	1	CR1_HUMAN
4	215	28.6	263	1	VCP_VACCV
5	199	26.5	547	1	C4BP_HUMAN
6	197.5	26.3	610	1	C4BP_BOVIN
7	182	24.2	558	1	C4BP_RAT
8	180.5	24.0	1231	1	CFAB_HUMAN
9	178.5	23.8	469	1	C4BP_MOUSE
10	178.5	23.8	1234	1	CFAB_MOUSE
11	167.5	22.3	390	1	DAF1_MOUSE
12	165	22.0	661	1	F13B_HUMAN
13	161	21.4	407	1	DAF2_MOUSE
14	159	21.2	377	1	MCP_HUMAN
15	158.5	21.1	440	1	DAF_PONDY
16	158.5	21.1	481	1	DAF_HUMAN
17	150	20.0	360	1	CYP_HSVSA
18	149	19.8	668	1	F13B_MOUSE
19	148	19.7	768	1	LEMA_MOUSE
20	147	19.6	330	1	PER1_HUMAN
21	146.5	19.5	768	1	LEMA_RAT
22	144	19.2	3019	1	LEP_CANE
23	142	18.9	1019	1	LEP_CACTH
24	141.5	18.8	507	1	DAF_CAVIO
25	141	18.8	252	1	C4BP_HUMAN
26	140	18.6	151	1	CFAB_PIG
27	140	18.6	258	1	C4BP_RAT
28	140	18.6	752	1	CR2_HUMAN
29	137	18.2	345	1	APDH_MOUSE
30	136.5	18.2	958	1	HUG_MOUSE
31	134.5	17.9	270	1	PHR2_HUMAN
32	134	17.8	297	1	APR_RAT
33	133	17.7	764	1	CFAB_HUMAN

34	132	17.6	345	1	APOLCANFA
35	131	17.4	202	1	APAR_PIG
36	131	17.4	685	1	CFAB_BOVIN
37	131	17.4	769	1	LEM2_MOUSE
38	129.5	17.2	549	1	LEM2_PIG
39	127	16.9	345	1	APOLBOVIN
40	127	16.9	345	1	APOLHUMAN
41	127	16.9	646	1	LEM3_BOVIN
42	123.5	16.4	612	1	LEM2_MOUSE
43	123	16.4	934	1	CO6_HUMAN
44	122	16.2	611	1	LEM2CANFA
45	122	16.2	761	1	CFAB_MOUSE
46	120.5	16.0	760	1	CO2_MOUSE
47	120	16.0	610	1	LEM2_HUMAN
48	119	15.8	331	1	PHR4_HUMAN
49	118	15.7	330	1	PHR4_HUMAN
50	118	15.7	484	1	LEM2_PIG
51	117.5	15.6	551	1	LEM2_RHAT
52	114	15.2	830	1	LEM3_HUMAN
53	110.5	14.7	843	1	CO7_HUMAN
54	110	14.6	485	1	LEM2_BOVIN
55	102.5	13.6	317	1	VR05_VACCV
56	102	13.6	464	1	SRPX_RAT
57	101.5	13.5	686	1	MAS2_HUMAN
58	99.5	13.2	317	1	VR05_VACCV
59	94	12.6	317	1	VR05_VACCV
60	94	12.5	464	1	SRPX_HUMAN
61	93	12.4	347	1	HPT_RAT
62	90.5	12.1	317	1	VR05_VACCV
63	86	11.5	699	1	GRAC_HUMAN
64	85.5	11.4	926	1	PER2_PIG
65	85	11.3	198	1	C4BP_BOVIN
66	85	11.3	347	1	HPT_MOUSE
67	84	11.2	370	1	LEM1_BOVIN
68	83.5	11.1	406	1	HPT2_HUMAN
69	83	11.1	250	1	CFBL_HUMAN
70	83	11.1	933	1	PERT_HUMAN
71	82.5	11.0	810	1	PLMN_MOUSE
72	81	10.8	372	1	LEM1_MOUSE
73	81	10.8	372	1	LEM1_PAPPA
74	80	10.7	372	1	LEM1_HUMAN
75	80	10.7	372	1	LEM1_PANIR
76	79	10.5	704	1	CRAR_MOUSE
77	78.5	10.5	1627	1	PAPA_HUMAN
78	78.5	10.5	3562	1	POCV_CHICK
79	77	10.3	372	1	LEM1_PONY
80	74	9.9	347	1	HPT_PIG
81	73.5	9.8	470	1	NRAM_CACTH
82	73.5	9.8	662	1	NR02_MOUSE
83	73.5	9.8	662	1	NR02_RAT
84	73.5	9.8	684	1	XYNA_CACTH
85	73.5	9.8	812	1	PLMN_BOVIN
86	72.5	9.7	378	1	CXAL_XENLA
87	72.5	9.7	470	1	NRAM_HUMAN
88	72.5	9.7	912	1	FCGB_BOVIN
89	71.5	9.5	710	1	NRAM_HUMAN
90	71.5	9.5	628	1	IO_HUMAN
91	71.5	9.5	688	1	CIS_HUMAN
92	71.5	9.5	705	1	CLIR_HUMAN
93	71.5	9.5	883	1	FCGB_MOUSE
94	71.5	9.5	883	1	FCGB_RAT
95	71	9.5	812	1	LEMN_MOUSE
96	71	9.5	1700	1	HAR3_HUMAN
97	71	9.5	2738	1	POCV_RAT
98	70.5	9.4	335	1	C4BP_MOUSE
99	70.5	9.4	662	1	MM02_RHAT
100	69.5	9.3	262	1	MUR1_HUMAN

ALL SUMMARIES

RESULT 1


```

FT CARBOHYD 682 682 N LINKED (GLUCNA... ) (POTENTIAL) .
FT CARBOHYD 800 800 N LINKED (GLUCNA... ) (POTENTIAL) .
FT CARBOHYD 823 823 N LINKED (GLUCNA... ) (POTENTIAL) .
FT CARBOHYD 861 861 N LINKED (GLUCNA... ) (POTENTIAL) .
FT CARBOHYD 911 911 N LINKED (GLUCNA... ) (POTENTIAL) .
FT VARSPLIC 499 524 MISSING (IN SHORT ISFORM) .
FT VARSPLIC 525 556 TPVYQWGTTHRELRFRSWHGHLM (IN SHORT
ISFORM) .
FT CONFLICT 667 667 Q - D (IN REF. 2) .
FT CONFLICT 902 902 Q - C (IN REF. 2) .
FT CONFLICT 906 906 H - L (IN REF. 2) .
SQ SEQUENCE 1033 AA: 112974 MW: 1749084A07847ADA CQC64;
Query Match 100.0%; Score 751; DB 1; Length 1033;
Seqs: total Similarity 100.0%; Pred. No. 3,20-68;
Matches 144; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GTSNCSPPPLNGRISYSIPVAGVIVIRYSCSGIEPLDGPKSLDGPVTKRWKVA 60
DB 20 GTSNCSPPPLNGRISYSIPVAGVIVIRYSCSGIEPLDGPKSLDGPVTKRWKVA 79
QY 61 TKEVFNKYSSTPEPVIAGVYKRWGSDYKRGDSVIFACKTINSMNKNSVWCGANNMWG 120
DB 80 PKCEVFNKYSSTPEPVIAGVYKRWGSDYKRGDSVIFACKTINSMNKNSVWCGANNMWG 139
QY 121 PTPGVSVPEPDE 134
DB 140 PTPGVSVPEPDE 153
RESULT 2
CR2_MOUSE
ID CR2_MOUSE STANDARD PRT: 1025 AA.
AC P19070;
DT 01 NOV 1990 (Rel. 16, Created)
DI 01 NOV 1990 (Rel. 16, Last sequence update)
DE 01 FEB 1996 (Rel. 33, Last annotation update)
DE Complement receptor type 2 precursor (Cr2) (Complement C3d receptor).
GN CR2.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
OX NBL TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BA1B7;
RX MEDLINE=90229735; PubMed 2139457;
RA Fingeroth J.D.;
RT "Comparative structure and evolution of murine CR2. The homolog of
the human C3d/ERV receptor (Cr2).";
RI J. Immunol. 144:3458-3467(1990).
RN [2]
RP SEQUENCE OF 12-1025 FROM N.A.
RX MEDLINE=91010789; PubMed=2145466;
RA Molina H., Kinoshita T., Inoue K., Carel J.C., Holers V.M.;
RT "A molecular and immunochemical characterization of mouse CR2.
Evidence for a single gene model of mouse complement receptors 1 and
2.";
RI J. Immunol. 145:2974-2983(1990).
RN [3]
RP SEQUENCE OF 43-401 AND 991-1025 FROM N.A.
RX MEDLINE=89098890; PubMed=2784485;
RA Fingeroth J.D., Houdiet M.A., Levy D.N., Strominger J.L.;
RT "Identification of murine complement receptor type 2.";
RI Proc. Natl. Acad. Sci. U.S.A. 86:242-246(1989).
RN [4]
RP SEQUENCE OF 289-1025 FROM N.A.
RX MEDLINE=89481350; PubMed=2528587;
RA Korte C.B., Paul M.S., Acquier M., Weiss J.J., Weiss J.H.;
RT "Murine complement receptor gene family. II. Identification and
characterization of the murine homolog (Cr2) to human CR2 and its
molecular linkage to Cr1.";

```

```

RL J. Immunol. 143:2058-2067(1989).
CC -1- FUNCTION: RECEPTOR FOR COMPLEMENT C3D. PARTICIPATES IN B
CC LYMPHOCYTES ACTIVATION.
CC -1- SUBCELLULAR LOCATION: Type I membrane protein.
CC -1- TISSUE SPECIFICITY: B LYMPHOCYTES.
CC -1- SIMILARITY: TO HUMAN C3D/ERV RECEPTOR (C021).
CC -1- SIMILARITY: CONTAINS 15 SUSHI (SCF) DOMAINS.
CC
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/faq-usa.html
CC or send an email to license@sib-sib.ch).
CC
EMBL: M81083: AAA37451.1;
EMBL: M35684: AAA37448.1;
EMBL: M61132: AAA63295.1;
EMBL: M35685: AAA37450.1; ALT_SEQ.
EMBL: M29281: AAA37447.1;
PIR: A43526: A43526.
HSSP: P10598: IVD.
MGD: MGI:88489: Cr2.
InterPro: IP000436; Sushi_SCR_CCP.
PIR: PFC084; sushi: 14.
SMART: SMC0042; CCP: 14.
KW Complement pathway; Glycoprotein; Transmembrane; Repeat; Signal;
Receptor; Sushi.
FT SIGNAL 1 11
FT CHAIN 12 1025 COMPLEMENT RECEPTOR TYPE 2.
FT DOMAIN 12 963 EXTRACELLULAR (POTENTIAL);
FT TRANSMEM 964 990 POTENTIAL...
FT DOMAIN 991 1025 CYTOPLASMIC (POTENTIAL).
FT DOMAIN 13 74 SUSHI 1
FT DOMAIN 81 139 SUSHI 2
FT DOMAIN 145 203 SUSHI 3
FT DOMAIN 206 264 SUSHI 4
FT DOMAIN 267 335 SUSHI 5
FT DOMAIN 342 399 SUSHI 6
FT DOMAIN 401 458 SUSHI 7
FT DOMAIN 462 515 SUSHI 8
FT DOMAIN 518 586 SUSHI 9
FT DOMAIN 593 649 SUSHI 10
FT DOMAIN 653 705 SUSHI 11
FT DOMAIN 708 769 SUSHI 12
FT DOMAIN 777 834 SUSHI 13
FT DOMAIN 840 898 SUSHI 14
FT DOMAIN 901 959 SUSHI 15
FT DISULFID 14 56 BY SIMILARITY.
FT DISULFID 42 73 HY SIMILARITY.
FT DISULFID 82 124 HY SIMILARITY.
FT DISULFID 110 138 HY SIMILARITY.
FT DISULFID 146 189 BY SIMILARITY.
FT DISULFID 175 202 BY SIMILARITY.
FT DISULFID 207 248 BY SIMILARITY.
FT DISULFID 234 263 HY SIMILARITY.
FT DISULFID 268 317 HY SIMILARITY.
FT DISULFID 297 334 HY SIMILARITY.
FT DISULFID 343 385 BY SIMILARITY.
FT DISULFID 371 398 BY SIMILARITY.
FT DISULFID 402 445 BY SIMILARITY.
FT DISULFID 431 458 HY SIMILARITY.
FT DISULFID 463 501 HY SIMILARITY.
FT DISULFID 487 514 HY SIMILARITY.
FT DISULFID 519 568 BY SIMILARITY.
FT DISULFID 548 585 BY SIMILARITY.
FT DISULFID 594 636 BY SIMILARITY.
FT DISULFID 622 649 HY SIMILARITY.
FT DISULFID 654 684 HY SIMILARITY.
FT DISULFID 675 704 HY SIMILARITY.
FT DISULFID 709 752 BY SIMILARITY.
FT DISULFID 738 769 BY SIMILARITY.

```

"Human C4b receptor (CRI). Demonstration of long homologies
 repeat in domains that are composed of the short consensus repeats
 characteristics of C3/C4 binding proteins." J. Exp. Med. 165:1095-1112(1987).
 [3]
 SEQUENCE OF 761-784; 841-845 AND 1179-1195 FROM N.A.
 MEDLINE-86067975; PubMed-2944745;
 Wong W.W., Kleinstein L.R., Smith J.A., Weis J.H., Pearson D.L.
 Identification of a partial cDNA clone for the human receptor for
 complement fragments C3b/C4b." J. Exp. Med. 165:1095-1112(1987).
 [3]
 FUNCTION: CRI, PRESENT ON ERYTHROCYTES, LEUKOCYTES, GLomerular
 cells, AND SPLENIC FOLLICULAR DENDRITIC CELLS. MEDIATES THE
 BINDING BY THESE CELLS OF PARTICLES AND IMMUNE COMPLEXES THAT HAVE
 ACTIVATED COMPLEMENT.
 [3]
 SUBUNIT: MONOMER.
 [3]
 SUBCELLULAR LOCATION: TYPE 1 membrane protein.
 [3]
 POLYMORPHISM: CRI IS RESPONSIBLE FOR THE KNOWN BLOOD GROUP SYSTEM
 [3]
 MISCELLANEOUS: SEVEN SHORT CONSENSUS REPEATS (SCR) CONSTITUTE A
 LONG HOMOLOGOUS REPEAT (LHR). THE N-TERMINAL TWO SCRS OF LHR A
 CONTAIN A SITE DETERMINING C4 SPECIFICITY, AND THE N-TERMINAL
 TWO SCRS OF LHR-B AND -C EACH HAVE A SITE DETERMINING C4
 SPECIFICITY.
 [3]
 MISCELLANEOUS: THIS IS THE SEQUENCE OF THE F ALL-TYPE OF CRI
 [3]
 SIMILARITY: CONTAINS 40 SUSHI (SCR) DOMAINS.
 [3]
 THIS SWISS PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Abuse by and for commercial
 entities related to a disease or chemical (see below) will be prosecuted
 or send an email to license@lsb.scrib.ch).
 [3]
 EMBL: M19569; AAA52297.1;
 EMBL: M19570; AAA52298.1;
 EMBL: M19571; AAA52299.1;
 EMBL: Y00816; CAA68755.1;
 EMBL: X05409; CAA28943.1;
 EMBL: A28547; A24507.
 EMBL: A24748; A24748.
 EMBL: B24748; B24748.
 EMBL: C24748; C24748.
 EMBL: S03843; S03843.
 EMBL: P08603; IHL1.
 EMBL: I20620.
 EMBL: IPR000436; Sushi_SCR_CRP.
 EMBL: P08604; sushi; 40.
 SMART: SM00042; GCP; 40.
 Complement pathway: Glycoprotein; Transmembrane; Repeat; Signal;
 Receptor; Sushi; Blood group antigen.
 SIGNAL: 1 41
 CHAIN: 42 2039
 DOMAIN: 42 1971
 TRANSMEM: 1972 1996
 DOMAIN: 1997 2039
 MOD_RES: 42 42
 DOMAIN: 42 100
 DOMAIN: 103 162
 DOMAIN: 165 233
 DOMAIN: 247 294
 DOMAIN: 296 354
 DOMAIN: 357 417
 DOMAIN: 420 488
 DOMAIN: 492 550
 DOMAIN: 553 612
 DOMAIN: 615 683
 DOMAIN: 687 744
 DOMAIN: 746 804
 DOMAIN: 807 867
 DOMAIN: 870 938
 DOMAIN: 942 1000
 DOMAIN: 1003 1062

FT	DOMAIN	1065	1143	SUSHI C4.	FT	DISULFID	1948	1965	HY SIM LAF TY.
FT	DOMAIN	1147	1194	SUSHI C4.	FT	CARBOHYD	55	56	N-LINKED GLCNAC (POTENTIAL)
FT	DOMAIN	1196	1244	SUSHI C5.	FT	CARBOHYD	212	252	N-LINKED GLCNAC (POTENTIAL)
FT	DOMAIN	1257	1317	SUSHI C6.	FT	CARBOHYD	410	410	N-LINKED GLCNAC (POTENTIAL)
FT	DOMAIN	1320	1408	SUSHI C7.	FT	CARBOHYD	447	447	N-LINKED GLCNAC (POTENTIAL)
FT	DOMAIN	1395	1453	SUSHI D1.	FT	CARBOHYD	509	509	N-LINKED GLCNAC (POTENTIAL)
FT	DOMAIN	1456	1515	SUSHI D2.	FT	CARBOHYD	578	578	N-LINKED GLCNAC (POTENTIAL)
FT	DOMAIN	1518	1586	SUSHI D3.	FT	CARBOHYD	702	702	N-LINKED GLCNAC (POTENTIAL)
FT	DOMAIN	1590	1647	SUSHI D4.	FT	CARBOHYD	860	860	N-LINKED GLCNAC (POTENTIAL)
FT	DOMAIN	1649	1707	SUSHI D5.	FT	CARBOHYD	897	897	N-LINKED GLCNAC (POTENTIAL)
FT	DOMAIN	1710	1770	SUSHI D6.	FT	CARBOHYD	959	959	N-LINKED GLCNAC (POTENTIAL)
FT	DOMAIN	1774	1841	SUSHI D7.	FT	CARBOHYD	1028	1028	N-LINKED GLCNAC (POTENTIAL)
FT	DOMAIN	1847	1905	SUSHI E1.	FT	CARBOHYD	1152	1152	N-LINKED GLCNAC (POTENTIAL)
FT	DOMAIN	1908	1966	SUSHI E2.	FT	CARBOHYD	1310	1310	N-LINKED GLCNAC (POTENTIAL)
FT	DISULFID	43	86	BY SIMILARITY.	FT	CARBOHYD	1481	1481	N-LINKED GLCNAC (POTENTIAL)
FT	DISULFID	74	99	BY SIMILARITY.	FT	CARBOHYD	1504	1504	N-LINKED GLCNAC (POTENTIAL)
FT	DISULFID	104	145	BY SIMILARITY.	FT	CARBOHYD	1534	1534	N-LINKED GLCNAC (POTENTIAL)
FT	DISULFID	111	161	BY SIMILARITY.	FT	CARBOHYD	1540	1540	N-LINKED GLCNAC (POTENTIAL)
FT	DISULFID	166	215	BY SIMILARITY.	FT	CARBOHYD	1605	1605	N-LINKED GLCNAC (POTENTIAL)
FT	DISULFID	195	242	BY SIMILARITY.	FT	CARBOHYD	1763	1763	N-LINKED GLCNAC (POTENTIAL)
FT	DISULFID	248	280	BY SIMILARITY.	FT	CARBOHYD	1908	1908	N-LINKED GLCNAC (POTENTIAL)
FT	DISULFID	285	294	BY SIMILARITY.	FT	CARBOHYD	2039	2039	N-LINKED GLCNAC (POTENTIAL)
FT	DISULFID	297	340	BY SIMILARITY.	FT	CARBOHYD	2039	2039	N-LINKED GLCNAC (POTENTIAL)
FT	DISULFID	325	353	BY SIMILARITY.	FT	CARBOHYD	2039	2039	N-LINKED GLCNAC (POTENTIAL)
FT	DISULFID	353	400	BY SIMILARITY.	FT	CARBOHYD	2039	2039	N-LINKED GLCNAC (POTENTIAL)
FT	DISULFID	386	416	BY SIMILARITY.	FT	CARBOHYD	2039	2039	N-LINKED GLCNAC (POTENTIAL)
FT	DISULFID	421	470	BY SIMILARITY.	FT	CARBOHYD	2039	2039	N-LINKED GLCNAC (POTENTIAL)
FT	DISULFID	450	487	BY SIMILARITY.	FT	CARBOHYD	2039	2039	N-LINKED GLCNAC (POTENTIAL)
FT	DISULFID	494	546	BY SIMILARITY.	FT	CARBOHYD	2039	2039	N-LINKED GLCNAC (POTENTIAL)
FT	DISULFID	524	569	BY SIMILARITY.	FT	CARBOHYD	2039	2039	N-LINKED GLCNAC (POTENTIAL)
FT	DISULFID	554	595	BY SIMILARITY.	FT	CARBOHYD	2039	2039	N-LINKED GLCNAC (POTENTIAL)
FT	DISULFID	581	631	BY SIMILARITY.	FT	CARBOHYD	2039	2039	N-LINKED GLCNAC (POTENTIAL)
FT	DISULFID	616	665	BY SIMILARITY.	FT	CARBOHYD	2039	2039	N-LINKED GLCNAC (POTENTIAL)
FT	DISULFID	645	682	BY SIMILARITY.	FT	CARBOHYD	2039	2039	N-LINKED GLCNAC (POTENTIAL)
FT	DISULFID	688	740	BY SIMILARITY.	FT	CARBOHYD	2039	2039	N-LINKED GLCNAC (POTENTIAL)
FT	DISULFID	716	743	BY SIMILARITY.	FT	CARBOHYD	2039	2039	N-LINKED GLCNAC (POTENTIAL)
FT	DISULFID	747	790	BY SIMILARITY.	FT	CARBOHYD	2039	2039	N-LINKED GLCNAC (POTENTIAL)
FT	DISULFID	776	803	BY SIMILARITY.	FT	CARBOHYD	2039	2039	N-LINKED GLCNAC (POTENTIAL)
FT	DISULFID	808	850	BY SIMILARITY.	FT	CARBOHYD	2039	2039	N-LINKED GLCNAC (POTENTIAL)
FT	DISULFID	846	866	BY SIMILARITY.	FT	CARBOHYD	2039	2039	N-LINKED GLCNAC (POTENTIAL)
FT	DISULFID	871	920	BY SIMILARITY.	FT	CARBOHYD	2039	2039	N-LINKED GLCNAC (POTENTIAL)
FT	DISULFID	900	947	BY SIMILARITY.	FT	CARBOHYD	2039	2039	N-LINKED GLCNAC (POTENTIAL)
FT	DISULFID	943	986	BY SIMILARITY.	FT	CARBOHYD	2039	2039	N-LINKED GLCNAC (POTENTIAL)
FT	DISULFID	973	999	BY SIMILARITY.	FT	CARBOHYD	2039	2039	N-LINKED GLCNAC (POTENTIAL)
FT	DISULFID	1004	1045	BY SIMILARITY.	FT	CARBOHYD	2039	2039	N-LINKED GLCNAC (POTENTIAL)
FT	DISULFID	1031	1061	BY SIMILARITY.	FT	CARBOHYD	2039	2039	N-LINKED GLCNAC (POTENTIAL)
FT	DISULFID	1066	1115	BY SIMILARITY.	FT	CARBOHYD	2039	2039	N-LINKED GLCNAC (POTENTIAL)
FT	DISULFID	1095	1142	BY SIMILARITY.	FT	CARBOHYD	2039	2039	N-LINKED GLCNAC (POTENTIAL)
FT	DISULFID	1148	1180	BY SIMILARITY.	FT	CARBOHYD	2039	2039	N-LINKED GLCNAC (POTENTIAL)
FT	DISULFID	1166	1193	BY SIMILARITY.	FT	CARBOHYD	2039	2039	N-LINKED GLCNAC (POTENTIAL)
FT	DISULFID	1197	1240	BY SIMILARITY.	FT	CARBOHYD	2039	2039	N-LINKED GLCNAC (POTENTIAL)
FT	DISULFID	1226	1253	BY SIMILARITY.	FT	CARBOHYD	2039	2039	N-LINKED GLCNAC (POTENTIAL)
FT	DISULFID	1258	1400	BY SIMILARITY.	FT	CARBOHYD	2039	2039	N-LINKED GLCNAC (POTENTIAL)
FT	DISULFID	1286	1316	BY SIMILARITY.	FT	CARBOHYD	2039	2039	N-LINKED GLCNAC (POTENTIAL)
FT	DISULFID	1321	1370	BY SIMILARITY.	FT	CARBOHYD	2039	2039	N-LINKED GLCNAC (POTENTIAL)
FT	DISULFID	1350	1387	BY SIMILARITY.	FT	CARBOHYD	2039	2039	N-LINKED GLCNAC (POTENTIAL)
FT	DISULFID	1396	1449	BY SIMILARITY.	FT	CARBOHYD	2039	2039	N-LINKED GLCNAC (POTENTIAL)
FT	DISULFID	1426	1452	BY SIMILARITY.	FT	CARBOHYD	2039	2039	N-LINKED GLCNAC (POTENTIAL)
FT	DISULFID	1457	1498	BY SIMILARITY.	FT	CARBOHYD	2039	2039	N-LINKED GLCNAC (POTENTIAL)
FT	DISULFID	1484	1504	BY SIMILARITY.	FT	CARBOHYD	2039	2039	N-LINKED GLCNAC (POTENTIAL)
FT	DISULFID	1519	1568	BY SIMILARITY.	FT	CARBOHYD	2039	2039	N-LINKED GLCNAC (POTENTIAL)
FT	DISULFID	1548	1585	BY SIMILARITY.	FT	CARBOHYD	2039	2039	N-LINKED GLCNAC (POTENTIAL)
FT	DISULFID	1591	1643	BY SIMILARITY.	FT	CARBOHYD	2039	2039	N-LINKED GLCNAC (POTENTIAL)
FT	DISULFID	1619	1646	BY SIMILARITY.	FT	CARBOHYD	2039	2039	N-LINKED GLCNAC (POTENTIAL)
FT	DISULFID	1650	1694	BY SIMILARITY.	FT	CARBOHYD	2039	2039	N-LINKED GLCNAC (POTENTIAL)
FT	DISULFID	1679	1706	BY SIMILARITY.	FT	CARBOHYD	2039	2039	N-LINKED GLCNAC (POTENTIAL)
FT	DISULFID	1711	1753	BY SIMILARITY.	FT	CARBOHYD	2039	2039	N-LINKED GLCNAC (POTENTIAL)
FT	DISULFID	1739	1769	BY SIMILARITY.	FT	CARBOHYD	2039	2039	N-LINKED GLCNAC (POTENTIAL)
FT	DISULFID	1774	1823	BY SIMILARITY.	FT	CARBOHYD	2039	2039	N-LINKED GLCNAC (POTENTIAL)
FT	DISULFID	1803	1840	BY SIMILARITY.	FT	CARBOHYD	2039	2039	N-LINKED GLCNAC (POTENTIAL)
FT	DISULFID	1848	1891	BY SIMILARITY.	FT	CARBOHYD	2039	2039	N-LINKED GLCNAC (POTENTIAL)
FT	DISULFID	1877	1904	BY SIMILARITY.	FT	CARBOHYD	2039	2039	N-LINKED GLCNAC (POTENTIAL)
FT	DISULFID	1909	1952	BY SIMILARITY.	FT	CARBOHYD	2039	2039	N-LINKED GLCNAC (POTENTIAL)

Query Match 29.2% Score 219 DB 1 Length 2039;

Best Local Similarity 36.7% Pred. No. 126-14;

Matches 51: Conservative 17: Mismatches 35: Indels 16: Gaps 6.

QY 2 ISGSPPTINRISYYS---TPIAVGVIRYS-----SCTERLIGKSLCTIKKV 52

DB 1517 ISCEPPTISNG--DPSNNRTSFNCTIVVYQTHGIDGQHPHGVGRSVCISKDQ 1574

QY 53 DCTWIKAKCEYFNKYSSCEPPIVQDYPIK231PVPEDSDVFAKINEMNKRVS 111

DB 1575 VGVWSSPEKISTNK---CTAPEVENAIRVGNKSTSEIIFKPEQPGFVWVGSHTV 1961

QY 112 WCCANNMWTPIKATCVSV 130

DB 1632 QCTINGWCP-KIPHCSPV 1649

RESULT 4

VCP_VACCV

ID VCP_VACCV

AC P10998

DT 01-JUL-1989 (rel. 11, Created)

DT 01-JUL-1989 (rel. 11, Last sequence update)

DT 16-OCT-2001 (rel. 40, Last annotation update)

DE Complement control protein precursor (v3r) (Secretory protein 45)

DE (protein C3) (28 kDa protein).

GN C3L

OS Vaccinia virus (strain WR), and

OS Vaccinia virus (strain Copenhagen).

OC Viruses; dsDNA viruses, no RNA stage; Poxviridae; Chordopoxvirinae;

OC Orthopoxvirus

GX NCBI_TaxID=10254, 10249;

RN [1]

RP SEQUENCE FROM N.A., AND SEQUENCE OF 40-47.

RC STRAIN-WP;

EX MEDLINE=88318974; PubMed=3412473;

RA Kotwal G.J., Moss B.;

RT "Vaccinia virus encodes a secretory polypeptide structurally related

to complement control proteins.";

RL Nature 335:176-178(1988).

RN [2]

RP SEQUENCE FROM N.A.

RC STRAIN-WP;

EX MEDLINE=89073956; PubMed=2849238;

RA Kotwal G.J., Moss B.;

RT "Analysis of a large cluster of non-essential genes deleted from a

vaccinia virus terminal transposition mutant.";

RL Virology 167:524-537(1988).

RN [3]

RP SEQUENCE FROM N.A.


```

FT SIGNAL 1 56
FT CHAIN 57 469
FT DOMAIN 57 116
FT DOMAIN 119 177
FT DOMAIN 180 241
FT DOMAIN 244 400
FT DOMAIN 402 456
FT DOMAIN 458 414
FT DISULFID 58 103
FT DISULFID 88 115
FT DISULFID 120 160
FT DISULFID 146 176
FT DISULFID 181 223
FT DISULFID 239 243
FT DISULFID 245 287
FT DISULFID 273 299
FT DISULFID 303 343
FT DISULFID 429 455
FT DISULFID 486 413
FT CARBOHYD 74 74
FT CARBOHYD 227 227
FT CARBOHYD 275 275
FT CARBOHYD 292 292
FT CARBOHYD 366 366
FT CARBOHYD 481 481
FT CARBOHYD 428 428
SQ SEQUENCE 469 AA; 51551 MW; 41E137C8B08C6321 CRC64;

Query Match 24.8%; Score 178.5; DB 1; Length 469;
Best Local Similarity 33.4%; Pred. No. 1e-10;
Matches 45; Conservative 14; Mismatches 54; Indels 23; Gaps 7.

QY 2 LSGSPDPTLNGRTS---YKSTPLAGVIVIVYKVSSTFRTIQKSKTGCTKIKKVGCTWD 57
DB 179 VKGGPPDPSNKKNSITELFY--FYNRG--ISYEDGCFKLVGSFFGKATVWNKIVFVWS 234

QY 58 KFAKCFENKYSKVPFP---IVPGYKTKGTSTPPPHPSVTFKATKTFNSMKNVSVW 112
DB 245 SSQPTCEK 10SQPNTHHGVIVSGYK---ATYTHRDSVRLACINGTVLGRHIV 286

QY 113 QANNWGPTRLPIC 127
DB 287 QANNW--SSLPIC 299

RESULT 10
CHAIN MOUSE
AC P06009. STANDARD; PRT: 1234 AA.
DI 01 JAN 1988 (rel. 06, Created)
DI 01 JAN 1988 (rel. 06, last sequence update)
DI 01 MAR 2002 (rel. 41, last annotation update)
DE Complement factor H precursor (Protein beta-1-H).
GN HFI or CFH
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
OX NCBI TaxID=10090;
PI 11
RX SEQUENCE FROM N.A.
RX MEDLINE=6624305; PubMed=2940596;
RA Kristensen F., Lack H.F.;
RI "Murine protein H is comprised of 20 repeating units, 61 amino acids
RI in length.*"
RL Proc. Natl. Acad. Sci. U.S.A. 84:3963-3967(1986).
RN 12.
RP SEQUENCE OF 119 FROM N.A.
RP STAIN-HAIR/CI;
RX MEDLINE=90148945; PubMed=2534512;
RA Minor changes P., Lack H.F., Vik B.P.;
RT Analysis of complement factor H mRNA expression: dexamethasone and
RT IFN-gamma increase the level of H in L cells.*

```

```

RL Biochemistry 28:9891-9897(1989).
RN 13.
RP SEQUENCE OF 116 FROM N.A.
RX MEDLINE=9011034; PubMed=2136885;
RA Natsuumi-Sakai S., Nonaka M., Nonaka Y.N., Shroffier D.J.,
RA Moriwaki K.;
RT "Demonstration of an unusual allelic variation of mouse factor H by
RT the complete cDNA sequence of the H 2 J101y10.*"
RL J. Immunol. 144:358-362(1990).
CC 1. FUNCTION: FACTOR H FUNCTIONS AS A COFACTOR IN THE INACTIVATION +
CC C3B BY FACTOR I AND ALSO INCREASES THE RATE OF DISSOCIATION OF THE
CC C3BBB COMPLEX (C3 CONVERTASE) AND THE (C4H)NRR COMPLEX (C5
CC CONVERTASE) IN THE ALTERNATIVE COMPLEMENT PATHWAY.
CC 2. POLYMORPHISM: TWO CODOMINANT ALLELES OF FACTOR H ARE PRESENT IN
CC MICE.
CC 3. SIMILARITY: CONTAINS 20 SUSHI (S-CK) DOMAINS.
CC 4. This SWISS-PORT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC at the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/).
CC or send an email to license@isb-sib.ch).
DR EMBL: M12660; AAA37759.1;
DR EMBL: J0289; AAA37795.1;
DR EMBL: M31974; AAA37762.1;
DR PIR: A26154; NMSSH.
DR HSSP: P08604; IHFI.
DR MGI: MGI:88485; Cfh.
DR InterPro: IPR000436; Sushi_SCR_CCP.
DR Pfam: PF00084; sushi; 20.
DR SMART: SM00042; CCP; 20.
KW Complement alternate pathway; Plasma; Glycoprotein; Repeat; Sushi;
KW Signal.
FT SIGNAL 1 18 BY SIMILARITY
FT CHAIN 19 1234 COMPLEMENT FACTOR H.
FT DOMAIN 20 81 SUSHI 1.
FT DOMAIN 84 142 SUSHI 2.
FT DOMAIN 145 206 SUSHI 3.
FT DOMAIN 209 263 SUSHI 4.
FT DOMAIN 266 321 SUSHI 5.
FT DOMAIN 324 386 SUSHI 6.
FT DOMAIN 388 443 SUSHI 7.
FT DOMAIN 447 506 SUSHI 8.
FT DOMAIN 508 565 SUSHI 9.
FT DOMAIN 568 623 SUSHI 10.
FT DOMAIN 628 684 SUSHI 11.
FT DOMAIN 689 744 SUSHI 12.
FT DOMAIN 751 803 SUSHI 13.
FT DOMAIN 807 862 SUSHI 14.
FT DOMAIN 866 932 SUSHI 15.
FT DOMAIN 935 990 SUSHI 16.
FT DOMAIN 993 1049 SUSHI 17.
FT DOMAIN 1052 1108 SUSHI 18.
FT DOMAIN 1113 1169 SUSHI 19.
FT DOMAIN 1171 1234 SUSHI 20.
FT DISULFID 21 66 BY SIMILARITY.
FT DISULFID 52 80 BY SIMILARITY.
FT DISULFID 85 129 BY SIMILARITY.
FT DISULFID 114 141 BY SIMILARITY.
FT DISULFID 145 192 BY SIMILARITY.
FT DISULFID 173 205 BY SIMILARITY.
FT DISULFID 213 251 BY SIMILARITY.
FT DISULFID 257 262 BY SIMILARITY.
FT DISULFID 267 309 BY SIMILARITY.
FT DISULFID 294 320 BY SIMILARITY.
FT DISULFID 325 374 BY SIMILARITY.
FT DISULFID 357 385 BY SIMILARITY.
FT DISULFID 389 431 BY SIMILARITY.
FT DISULFID 415 442 BY SIMILARITY.
FT DISULFID 443 494 BY SIMILARITY.

```


ACTIVATION AT THE FETO PLACENTAL INTERFACE ON THE
SYNCYTIOTROPHBLAST LAYER OF PLACENTA.
1- SUBCELLULAR LOCATION: Type 1 membrane protein.
1- ALTERNATIVE PRODUCTS: AT LEAST 14 FORMS ARE PRODUCED BY
ALTERNATIVE SPLICING. THE FORM SHOWN IS THE TRANSCRIPT A. THE
GAMMA ISOFORM (TRANSCRIPTS A AND B) IS PREFERENTIALLY EXPRESSED IN
EMBRYO CELLS AND TERATOMA TUMORS. THE 66 KDa ALPHA ISOFORM AND THE
56 KDa BETA ISOFORM FOUND IN ALL TISSUES EXCEPT SPERM CORRESPOND
TO THE TRANSCRIPTS C AND D. AND THE TRANSCRIPTS E AND F
RESPECTIVELY. THE EXON 9 IS SPECIFICALLY DELETED IN SOME
PLACENTAE. THE EPSILON ISOFORM CORRESPONDS TO THE TRANSCRIPTS I
AND J. SPERMATOZOIA DELTA ISOFORM (TRANSCRIPT N) DELETE BOTH EXONS
12 AND 13. ALL TISSUES DIFFERENTIALLY SPLICED EXON 13
1- TISSUE SPECIFICITY: EXPRESSED BY ALMOST ALL CELLS EXCEPT
ERYTHROCYTES AND SOME BONE MARROW CELLS.
1- PTM: EXTENSIVELY O-GLYCOSYLATED IN THE SER/THIR-RICH DOMAIN
(PROBABLE).
1- SIMILARITY: CONTAINS 4 SUSHI (SCR) DOMAINS.
1- DATABASES: NAME-PROW, NCBI-CD guide CD46 entry;
WWW="http://www.ncbi.nlm.nih.gov/prow/cd/cd46.htm".

THIS SWISS-PROT ENTRY IS COPYRIGHT. IT IS PRODUCED THROUGH A COLLABORATION
BETWEEN THE SWISS INSTITUTE OF BIOINFORMATICS AND THE EMBL OUTSTATION -
THE EUROPEAN BIOINFORMATICS INSTITUTE. THERE ARE NO RESTRICTIONS ON ITS
USE BY NON PROFIT INSTITUTIONS AS LONG AS ITS CONTENT IS IN NO WAY
MODIFIED AND THIS STATEMENT IS NOT REMOVED. USAGE BY AND FOR COMMERCIAL
ENTITIES REQUIRES A LICENSE AGREEMENT (SEE <http://www.isb.sib.ch/announce/>)
OR SEND AN EMAIL TO license@isb.sib.ch.

EMBL: Y00651; CAA60675.1;
EMBL: S51940; AA24802.1;
EMBL: M58050; AA62843.1;
EMBL: A18585; CAA01400.1;
EMBL: S65879; AA013968.1;
FIR: S01896; S01896.
RSD: P10998; IVD0.
MIM: 129920.
InterPro: IPR000436; Sushi_SCR_2/P.
Pfam: PF00084; SUSHI_4.
SMART: SM00042; CCP; 4.
KW Complement pathway; Glycoprotein; Transmembrane; Repeat; Signal;
K1 Sushi; Alternative splicing.
F1 SIGNAL 1 44
F1 CHAIN 35 377 MEMBRANE COFACTOR PROTEIN.
F1 DOMAIN 45 328 EXTRACELLULAR (POTENTIAL).
F1 TRANSMEM 429 351 POTENTIAL.
F1 DOMAIN 352 377 CYTOPLASMIC (POTENTIAL).
F1 SUSHI 1 35
F1 SUSHI 2 98
F1 SUSHI 3 161
F1 SUSHI 4 227
F1 SUSHI 5 287
F1 SUSHI 6 311
F1 SUSHI 7 351
F1 SUSHI 8 351
F1 SUSHI 9 351
F1 SUSHI 10 351
F1 SUSHI 11 351
F1 SUSHI 12 351
F1 SUSHI 13 351
F1 SUSHI 14 351
F1 SUSHI 15 351
F1 SUSHI 16 351
F1 SUSHI 17 351
F1 SUSHI 18 351
F1 SUSHI 19 351
F1 SUSHI 20 351
F1 SUSHI 21 351
F1 SUSHI 22 351
F1 SUSHI 23 351
F1 SUSHI 24 351
F1 SUSHI 25 351
F1 SUSHI 26 351
F1 SUSHI 27 351
F1 SUSHI 28 351
F1 SUSHI 29 351
F1 SUSHI 30 351
F1 SUSHI 31 351
F1 SUSHI 32 351
F1 SUSHI 33 351
F1 SUSHI 34 351
F1 SUSHI 35 351
F1 SUSHI 36 351
F1 SUSHI 37 351
F1 SUSHI 38 351
F1 SUSHI 39 351
F1 SUSHI 40 351
F1 SUSHI 41 351
F1 SUSHI 42 351
F1 SUSHI 43 351
F1 SUSHI 44 351
F1 SUSHI 45 351
F1 SUSHI 46 351
F1 SUSHI 47 351
F1 SUSHI 48 351
F1 SUSHI 49 351
F1 SUSHI 50 351
F1 SUSHI 51 351
F1 SUSHI 52 351
F1 SUSHI 53 351
F1 SUSHI 54 351
F1 SUSHI 55 351
F1 SUSHI 56 351
F1 SUSHI 57 351
F1 SUSHI 58 351
F1 SUSHI 59 351
F1 SUSHI 60 351
F1 SUSHI 61 351
F1 SUSHI 62 351
F1 SUSHI 63 351
F1 SUSHI 64 351
F1 SUSHI 65 351
F1 SUSHI 66 351
F1 SUSHI 67 351
F1 SUSHI 68 351
F1 SUSHI 69 351
F1 SUSHI 70 351
F1 SUSHI 71 351
F1 SUSHI 72 351
F1 SUSHI 73 351
F1 SUSHI 74 351
F1 SUSHI 75 351
F1 SUSHI 76 351
F1 SUSHI 77 351
F1 SUSHI 78 351
F1 SUSHI 79 351
F1 SUSHI 80 351
F1 SUSHI 81 351
F1 SUSHI 82 351
F1 SUSHI 83 351
F1 SUSHI 84 351
F1 SUSHI 85 351
F1 SUSHI 86 351
F1 SUSHI 87 351
F1 SUSHI 88 351
F1 SUSHI 89 351
F1 SUSHI 90 351
F1 SUSHI 91 351
F1 SUSHI 92 351
F1 SUSHI 93 351
F1 SUSHI 94 351
F1 SUSHI 95 351
F1 SUSHI 96 351
F1 SUSHI 97 351
F1 SUSHI 98 351
F1 SUSHI 99 351
F1 SUSHI 100 351
F1 SUSHI 101 351
F1 SUSHI 102 351
F1 SUSHI 103 351
F1 SUSHI 104 351
F1 SUSHI 105 351
F1 SUSHI 106 351
F1 SUSHI 107 351
F1 SUSHI 108 351
F1 SUSHI 109 351
F1 SUSHI 110 351
F1 SUSHI 111 351
F1 SUSHI 112 351
F1 SUSHI 113 351
F1 SUSHI 114 351
F1 SUSHI 115 351
F1 SUSHI 116 351
F1 SUSHI 117 351
F1 SUSHI 118 351
F1 SUSHI 119 351
F1 SUSHI 120 351
F1 SUSHI 121 351
F1 SUSHI 122 351
F1 SUSHI 123 351
F1 SUSHI 124 351
F1 SUSHI 125 351
F1 SUSHI 126 351
F1 SUSHI 127 351
F1 SUSHI 128 351
F1 SUSHI 129 351
F1 SUSHI 130 351
F1 SUSHI 131 351
F1 SUSHI 132 351
F1 SUSHI 133 351
F1 SUSHI 134 351
F1 SUSHI 135 351
F1 SUSHI 136 351
F1 SUSHI 137 351
F1 SUSHI 138 351
F1 SUSHI 139 351
F1 SUSHI 140 351
F1 SUSHI 141 351
F1 SUSHI 142 351
F1 SUSHI 143 351
F1 SUSHI 144 351
F1 SUSHI 145 351
F1 SUSHI 146 351
F1 SUSHI 147 351
F1 SUSHI 148 351
F1 SUSHI 149 351
F1 SUSHI 150 351
F1 SUSHI 151 351
F1 SUSHI 152 351
F1 SUSHI 153 351
F1 SUSHI 154 351
F1 SUSHI 155 351
F1 SUSHI 156 351
F1 SUSHI 157 351
F1 SUSHI 158 351
F1 SUSHI 159 351
F1 SUSHI 160 351
F1 SUSHI 161 351
F1 SUSHI 162 351
F1 SUSHI 163 351
F1 SUSHI 164 351
F1 SUSHI 165 351
F1 SUSHI 166 351
F1 SUSHI 167 351
F1 SUSHI 168 351
F1 SUSHI 169 351
F1 SUSHI 170 351
F1 SUSHI 171 351
F1 SUSHI 172 351
F1 SUSHI 173 351
F1 SUSHI 174 351
F1 SUSHI 175 351
F1 SUSHI 176 351
F1 SUSHI 177 351
F1 SUSHI 178 351
F1 SUSHI 179 351
F1 SUSHI 180 351
F1 SUSHI 181 351
F1 SUSHI 182 351
F1 SUSHI 183 351
F1 SUSHI 184 351
F1 SUSHI 185 351
F1 SUSHI 186 351
F1 SUSHI 187 351
F1 SUSHI 188 351
F1 SUSHI 189 351
F1 SUSHI 190 351
F1 SUSHI 191 351
F1 SUSHI 192 351
F1 SUSHI 193 351
F1 SUSHI 194 351
F1 SUSHI 195 351
F1 SUSHI 196 351
F1 SUSHI 197 351
F1 SUSHI 198 351
F1 SUSHI 199 351
F1 SUSHI 200 351
F1 SUSHI 201 351
F1 SUSHI 202 351
F1 SUSHI 203 351
F1 SUSHI 204 351
F1 SUSHI 205 351
F1 SUSHI 206 351
F1 SUSHI 207 351
F1 SUSHI 208 351
F1 SUSHI 209 351
F1 SUSHI 210 351
F1 SUSHI 211 351
F1 SUSHI 212 351
F1 SUSHI 213 351
F1 SUSHI 214 351
F1 SUSHI 215 351
F1 SUSHI 216 351
F1 SUSHI 217 351
F1 SUSHI 218 351
F1 SUSHI 219 351
F1 SUSHI 220 351
F1 SUSHI 221 351
F1 SUSHI 222 351
F1 SUSHI 223 351
F1 SUSHI 224 351
F1 SUSHI 225 351
F1 SUSHI 226 351
F1 SUSHI 227 351
F1 SUSHI 228 351
F1 SUSHI 229 351
F1 SUSHI 230 351
F1 SUSHI 231 351
F1 SUSHI 232 351
F1 SUSHI 233 351
F1 SUSHI 234 351
F1 SUSHI 235 351
F1 SUSHI 236 351
F1 SUSHI 237 351
F1 SUSHI 238 351
F1 SUSHI 239 351
F1 SUSHI 240 351
F1 SUSHI 241 351
F1 SUSHI 242 351
F1 SUSHI 243 351
F1 SUSHI 244 351
F1 SUSHI 245 351
F1 SUSHI 246 351
F1 SUSHI 247 351
F1 SUSHI 248 351
F1 SUSHI 249 351
F1 SUSHI 250 351
F1 SUSHI 251 351
F1 SUSHI 252 351
F1 SUSHI 253 351
F1 SUSHI 254 351
F1 SUSHI 255 351
F1 SUSHI 256 351
F1 SUSHI 257 351
F1 SUSHI 258 351
F1 SUSHI 259 351
F1 SUSHI 260 351
F1 SUSHI 261 351
F1 SUSHI 262 351
F1 SUSHI 263 351
F1 SUSHI 264 351
F1 SUSHI 265 351
F1 SUSHI 266 351
F1 SUSHI 267 351
F1 SUSHI 268 351
F1 SUSHI 269 351
F1 SUSHI 270 351
F1 SUSHI 271 351
F1 SUSHI 272 351
F1 SUSHI 273 351
F1 SUSHI 274 351
F1 SUSHI 275 351
F1 SUSHI 276 351
F1 SUSHI 277 351
F1 SUSHI 278 351
F1 SUSHI 279 351
F1 SUSHI 280 351
F1 SUSHI 281 351
F1 SUSHI 282 351
F1 SUSHI 283 351
F1 SUSHI 284 351
F1 SUSHI 285 351
F1 SUSHI 286 351
F1 SUSHI 287 351
F1 SUSHI 288 351
F1 SUSHI 289 351
F1 SUSHI 290 351
F1 SUSHI 291 351
F1 SUSHI 292 351
F1 SUSHI 293 351
F1 SUSHI 294 351
F1 SUSHI 295 351
F1 SUSHI 296 351
F1 SUSHI 297 351
F1 SUSHI 298 351
F1 SUSHI 299 351
F1 SUSHI 300 351
F1 SUSHI 301 351
F1 SUSHI 302 351
F1 SUSHI 303 351
F1 SUSHI 304 351
F1 SUSHI 305 351
F1 SUSHI 306 351
F1 SUSHI 307 351
F1 SUSHI 308 351
F1 SUSHI 309 351
F1 SUSHI 310 351
F1 SUSHI 311 351
F1 SUSHI 312 351
F1 SUSHI 313 351
F1 SUSHI 314 351
F1 SUSHI 315 351
F1 SUSHI 316 351
F1 SUSHI 317 351
F1 SUSHI 318 351
F1 SUSHI 319 351
F1 SUSHI 320 351
F1 SUSHI 321 351
F1 SUSHI 322 351
F1 SUSHI 323 351
F1 SUSHI 324 351
F1 SUSHI 325 351
F1 SUSHI 326 351
F1 SUSHI 327 351
F1 SUSHI 328 351
F1 SUSHI 329 351
F1 SUSHI 330 351
F1 SUSHI 331 351
F1 SUSHI 332 351
F1 SUSHI 333 351
F1 SUSHI 334 351
F1 SUSHI 335 351
F1 SUSHI 336 351
F1 SUSHI 337 351
F1 SUSHI 338 351
F1 SUSHI 339 351
F1 SUSHI 340 351
F1 SUSHI 341 351
F1 SUSHI 342 351
F1 SUSHI 343 351
F1 SUSHI 344 351
F1 SUSHI 345 351
F1 SUSHI 346 351
F1 SUSHI 347 351
F1 SUSHI 348 351
F1 SUSHI 349 351
F1 SUSHI 350 351
F1 SUSHI 351 351
F1 SUSHI 352 351
F1 SUSHI 353 351
F1 SUSHI 354 351
F1 SUSHI 355 351
F1 SUSHI 356 351
F1 SUSHI 357 351
F1 SUSHI 358 351
F1 SUSHI 359 351
F1 SUSHI 360 351
F1 SUSHI 361 351
F1 SUSHI 362 351
F1 SUSHI 363 351
F1 SUSHI 364 351
F1 SUSHI 365 351
F1 SUSHI 366 351
F1 SUSHI 367 351
F1 SUSHI 368 351
F1 SUSHI 369 351
F1 SUSHI 370 351
F1 SUSHI 371 351
F1 SUSHI 372 351
F1 SUSHI 373 351
F1 SUSHI 374 351
F1 SUSHI 375 351
F1 SUSHI 376 351
F1 SUSHI 377 351
F1 SUSHI 378 351
F1 SUSHI 379 351
F1 SUSHI 380 351
F1 SUSHI 381 351
F1 SUSHI 382 351
F1 SUSHI 383 351
F1 SUSHI 384 351
F1 SUSHI 385 351
F1 SUSHI 386 351
F1 SUSHI 387 351
F1 SUSHI 388 351
F1 SUSHI 389 351
F1 SUSHI 390 351
F1 SUSHI 391 351
F1 SUSHI 392 351
F1 SUSHI 393 351
F1 SUSHI 394 351
F1 SUSHI 395 351
F1 SUSHI 396 351
F1 SUSHI 397 351
F1 SUSHI 398 351
F1 SUSHI 399 351
F1 SUSHI 400 351
F1 SUSHI 401 351
F1 SUSHI 402 351
F1 SUSHI 403 351
F1 SUSHI 404 351
F1 SUSHI 405 351
F1 SUSHI 406 351
F1 SUSHI 407 351
F1 SUSHI 408 351
F1 SUSHI 409 351
F1 SUSHI 410 351
F1 SUSHI 411 351
F1 SUSHI 412 351
F1 SUSHI 413 351
F1 SUSHI 414 351
F1 SUSHI 415 351
F1 SUSHI 416 351
F1 SUSHI 417 351
F1 SUSHI 418 351
F1 SUSHI 419 351
F1 SUSHI 420 351
F1 SUSHI 421 351
F1 SUSHI 422 351
F1 SUSHI 423 351
F1 SUSHI 424 351
F1 SUSHI 425 351
F1 SUSHI 426 351
F1 SUSHI 427 351
F1 SUSHI 428 351
F1 SUSHI 429 351
F1 SUSHI 430 351
F1 SUSHI 431 351
F1 SUSHI 432 351
F1 SUSHI 433 351
F1 SUSHI 434 351
F1 SUSHI 435 351
F1 SUSHI 436 351
F1 SUSHI 437 351
F1 SUSHI 438 351
F1 SUSHI 439 351
F1 SUSHI 440 351
F1 SUSHI 441 351
F1 SUSHI 442 351
F1 SUSHI 443 351
F1 SUSHI 444 351
F1 SUSHI 445 351
F1 SUSHI 446 351
F1 SUSHI 447 351
F1 SUSHI 448 351
F1 SUSHI 449 351
F1 SUSHI 450 351
F1 SUSHI 451 351
F1 SUSHI 452 351
F1 SUSHI 453 351
F1 SUSHI 454 351
F1 SUSHI 455 351
F1 SUSHI 456 351
F1 SUSHI 457 351
F1 SUSHI 458 351
F1 SUSHI 459 351
F1 SUSHI 460 351
F1 SUSHI 461 351
F1 SUSHI 462 351
F1 SUSHI 463 351
F1 SUSHI 464 351
F1 SUSHI 465 351
F1 SUSHI 466 351
F1 SUSHI 467 351
F1 SUSHI 468 351
F1 SUSHI 469 351
F1 SUSHI 470 351
F1 SUSHI 471 351
F1 SUSHI 472 351
F1 SUSHI 473 351
F1 SUSHI 474 351
F1 SUSHI 475 351
F1 SUSHI 476 351
F1 SUSHI 477 351
F1 SUSHI 478 351
F1 SUSHI 479 351
F1 SUSHI 480 351
F1 SUSHI 481 351
F1 SUSHI 482 351
F1 SUSHI 483 351
F1 SUSHI 484 351
F1 SUSHI 485 351
F1 SUSHI 486 351
F1 SUSHI 487 351
F1 SUSHI 488 351
F1 SUSHI 489 351
F1 SUSHI 490 351
F1 SUSHI 491 351
F1 SUSHI 492 351
F1 SUSHI 493 351
F1 SUSHI 494 351
F1 SUSHI 495 351
F1 SUSHI 496 351
F1 SUSHI 497 351
F1 SUSHI 498 351
F1 SUSHI 499 351
F1 SUSHI 500 351
F1 SUSHI 501 351
F1 SUSHI 502 351
F1 SUSHI 503 351
F1 SUSHI 504 351
F1 SUSHI 505 351
F1 SUSHI 506 351
F1 SUSHI 507 351
F1 SUSHI 508 351
F1 SUSHI 509 351
F1 SUSHI 510 351
F1 SUSHI 511 351
F1 SUSHI 512 351
F1 SUSHI 513 351
F1 SUSHI 514 351
F1 SUSHI 515 351
F1 SUSHI 516 351
F1 SUSHI 517 351
F1 SUSHI 518 351
F1 SUSHI 519 351
F1 SUSHI 520 351
F1 SUSHI 521 351
F1 SUSHI 522 351
F1 SUSHI 523 351
F1 SUSHI 524 351
F1 SUSHI 525 351
F1 SUSHI 526 351
F1 SUSHI 527 351
F1 SUSHI 528 351
F1 SUSHI 529 351
F1 SUSHI 530 351
F1 SUSHI 531 351
F1 SUSHI 532 351
F1 SUSHI 533 351
F1 SUSHI 534 351
F1 SUSHI 535 351
F1 SUSHI 536 351
F1 SUSHI 537 351
F1 SUSHI 538 351
F1 SUSHI 539 351
F1 SUSHI 540 351
F1 SUSHI 541 351
F1 SUSHI 542 351
F1 SUSHI 543 351
F1 SUSHI 544 351
F1 SUSHI 545 351
F1 SUSHI 546 351
F1 SUSHI 547 351
F1 SUSHI 548 351
F1 SUSHI 549 351
F1 SUSHI 550 351
F1 SUSHI 551 351
F1 SUSHI 552 351
F1 SUSHI 553 351
F1 SUSHI 554 351
F1 SUSHI 555 351
F1 SUSHI 556 351
F1 SUSHI 557 351
F1 SUSHI 558 351
F1 SUSHI 559 351
F1 SUSHI 560 351
F1 SUSHI 561 351
F1 SUSHI 562 351
F1 SUSHI 563 351
F1 SUSHI 564 351
F1 SUSHI 565 351
F1 SUSHI 566 351
F1 SUSHI 567 351
F1 SUSHI 568 351
F1 SUSHI 569 351
F1 SUSHI 570 351
F1 SUSHI 571 351
F1 SUSHI 572 351
F1 SUSHI 573 351
F1 SUSHI 574 351
F1 SUSHI 575 351
F1 SUSHI 576 351
F1 SUSHI 577 351
F1 SUSHI 578 351
F1 SUSHI 579 351
F1 SUSHI 580 351
F1 SUSHI 581 351
F1 SUSHI 582 351
F1 SUSHI 583 351
F1 SUSHI 584 351
F1 SUSHI 585 351
F1 SUSHI 586 351
F1 SUSHI 587 351
F1 SUSHI 588 351
F1 SUSHI 589 351
F1 SUSHI 590 351
F1 SUSHI 591 351
F1 SUSHI 592 351
F1 SUSHI 593 351
F1 SUSHI 594 351
F1 SUSHI 595 351
F1 SUSHI 596 351
F1 SUSHI 597 351
F1 SUSHI 598 351
F1 SUSHI 599 351
F1 SUSHI 600 351
F1 SUSHI 601 351
F1 SUSHI 602 351
F1 SUSHI 603 351
F1 SUSHI 604 351
F1 SUSHI 605 351
F1 SUSHI 606 351
F1 SUSHI 607 351
F1 SUSHI 608 351
F1 SUSHI 609 351
F1 SUSHI 610 351
F1 SUSHI 611 351
F1 SUSHI 612 351
F1 SUSHI 613 351
F1 SUSHI 614 351
F1 SUSHI 615 351
F1 SUSHI 616 351
F1 SUSHI 617 351
F1 SUSHI 618 351
F1 SUSHI 619 351
F1 SUSHI 620 351
F1 SUSHI 621 351
F1 SUSHI 622 351
F1 SUSHI 623 351
F1 SUSHI 624 351
F1 SUSHI 625 351
F1 SUSHI 626 351
F1 SUSHI 627 351
F1 SUSHI 628 351
F1 SUSHI 629 351
F1 SUSHI 630 351
F1 SUSHI 631 351
F1 SUSHI 632 351
F1 SUSHI 633 351
F1 SUSHI 634 351
F1 SUSHI 635 351
F1 SUSHI 636 351
F1 SUSHI 637 351
F1 SUSHI 638 351
F1 SUSHI 639 351
F1 SUSHI 640 351
F1 SUSHI 641 351
F1 SUSHI 642 351
F1 SUSHI 643 351
F1 SUSHI 644 351
F1 SUSHI 645 351
F1 SUSHI 646 351
F1 SUSHI 647 351
F1 SUSHI 648 351
F1 SUSHI 649 351
F1 SUSHI 650 351
F1 SUSHI 651 351
F1 SUSHI 652 351
F1 SUSHI 653 351
F1 SUSHI 654 351
F1 SUSHI 655 351
F1 SUSHI 656 351
F1 SUSHI 657 351
F1 SUSHI 658 351
F1 SUSHI 659 351
F1 SUSHI 660 351
F1 SUSHI 661 351
F1 SUSHI 662 351
F1 SUSHI 663 351
F1 SUSHI 664 351
F1 SUSHI 665 351
F1 SUSHI 666 351
F1 SUSHI 667 351
F1 SUSHI 668 351
F1 SUSHI 669 351
F1 SUSHI 670 351
F1 SUSHI 671 351
F1 SUSHI 672 351
F1 SUSHI 673 351
F1 SUSHI 674 351
F1 SUSHI 675 351
F1 SUSHI 676 351
F1 SUSHI 677 351
F1 SUSHI 678 351
F1 SUSHI 679 351
F1 SUSHI 680 351
F1 SUSHI 681 351
F1 SUSHI 682 351
F1 SUSHI 683 351
F1 SUSHI 684 351
F1 SUSHI 685 351
F1 SUSHI 686 351
F1 SUSHI 687 351
F1 SUSHI 688 351
F1 SUSHI 689 351
F1 SUSHI 690 351
F1 SUSHI 691 351
F1 SUSHI 692 351
F1 SUSHI 693 351
F1 SUSHI 694 351
F1 SUSHI 695 351
F1 SUSHI 696 351
F1 SUSHI 697 351
F1 SUSHI 698 351
F1 SUSHI 699 351
F1 SUSHI 700 351
F1 SUSHI 701 351
F1 SUSHI 702 351
F1 SUSHI 703 351
F1 SUSHI 704 351
F1 SUSHI 705 351
F1 SUSHI 706 351
F1 SUSHI 707 351
F1 SUSHI 708 351
F1 SUSHI 709 351
F1 SUSHI 710 351
F1 SUSHI 711 351
F1 SUSHI 712 351
F1 SUSHI 713 351
F1 SUSHI 714 351
F1 SUSHI 715 351
F1 SUSHI 716 351
F1 SUSHI 717 351
F1 SUSHI 718 351
F1 SUSHI 719 351
F1 SUSHI 720 351
F1 SUSHI 721 351
F1 SUSHI 722 351
F1 SUSHI 723 351
F1 SUSHI 724 351
F1 SUSHI 725 351
F1 SUSHI 726 351
F1 SUSHI 727 351
F1 SUSHI 728 351
F1 SUSHI 729 351
F1 SUSHI 730 351
F1 SUSHI 731 351
F1 SUSHI 732 351
F1 SUSHI 733 351
F1 SUSHI 734 351
F1 SUSHI 735 351
F1 SUSHI 736 351
F1 SUSHI 737 351
F1 SUSHI 738 351
F1 SUSHI 739 351
F1 SUSHI 740 351
F1 SUSHI 741 351
F1 SUSHI 742 351
F1 SUSHI 743 351
F1 SUSHI 744 351
F1 SUSHI 745 351
F1 SUSHI 746 351
F1 SUSHI 747 351
F1 SUSHI 748 351
F1 SUSHI 749 351
F1 SUSHI 750 351
F1 SUSHI 751 351
F1 SUSHI 752 351
F1 SUSHI 753 351
F1 SUSHI 754 351
F1 SUSHI 755 351
F1 SUSHI 756 351
F1 SUSHI

U00001	SMART: SMO00012; CYP: 4.																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																	</
--------	--------------------------	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	----

1 DOMAIN: THE FIRST SUSHI DOMAIN (SCR1) IS NOT NECESSARY FOR
 FUNCTION. SCR2 AND SCR4 PROVIDE THE PROPER CONFORMATION FOR THE
 ACTIVE SITE ON SCR3 (BY SIMILARITY).
 2 PDM: THE SER/TUR RICH DOMAIN IS HEAVILY O-GLYCOSYLATED.
 3 POLYMORPHISM: DAF IS RESPONSIBLE FOR THE CHROMER BLOOD GROUP
 SYSTEM. IT CONSISTS OF AT LEAST SEVEN HIGH-INCIDENCE (CR(A),
 TC(A), DR(A), ES(A), WES(B), DM* AND TP*) AND LOW INCIDENCE
 (TC(B), TC(C), AND WES(A)) ANTIGENS THAT RESIDE ON DAF. IN THE
 CHROMER PHENOTYPES DR(A) AND INAH THERE IS REDUCED OR ABSENT
 EXPRESSION OF DAF, RESPECTIVELY. IN THE CASE OF THE DR(A-)
 PHENOTYPE, A SINGLE NUCLEOTIDE SUBSTITUTION WITHIN EXON 5 ACCOUNTS
 FOR TWO CHANGES: A SIMPLE AMINO ACID SUBSTITUTION THAT IS THE
 BASIS OF THE ANTIGENIC VARIATION, AND AN ALTERNATIVE SPLICING
 EVENT THAT UNDERLIES THE DECREASED EXPRESSION OF DAF IN THIS
 PHENOTYPE.
 4 SIMILARITY: CONTAINS 4 SUSHI (SCR) DOMAINS.
 5 SIMILARITY: BELONGS TO THE REPEATS OF COMPLEMENT ACTIVATION
 (RCA) FAMILY.
 6 DATABASE: NAME: PROW. NOTE: CD guide CD55 entry:
 WWW="http://www.ncbi.nlm.nih.gov/prow/cd/cd55.htm".
 7 THIS SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL Outstation.
 8 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (see <http://www.isb-sib.ch/announcement/>
 or send an email to license@sib-sib.ch).
 9
 10 EMBL: M49516; AAA52169.1;
 11 EMBL: M40142; AAA52168.1;
 12 EMBL: BC001288; AA01288.1;
 13 EMBL: M15799; AAA52167.1;
 14 EMBL: U08576; AA48622.1;
 15 EMBL: M64653; AAA52170.1;
 16 EMBL: M64456; AAA52170.1; JOINED.
 17 EMBL: S72858; AAC60633.1;
 18 EMBL: B26459; B26459.
 19 EMBL: A26459; A26459.
 20 EMBL: S16187; S16187.
 21 EMBL: A19101; A19101.
 22 EMBL: S24148; S24148.
 23 EMBL: P08693; I08C.
 24 EMBL: L25240;
 25 InterPro: IPR000446; Sush1_scr_cpf.
 26 Pfam: PF00084; Sush1; 4.
 27 SMART: SM00042; ccp; 4.
 28 Complement pathway; Plasma; Glycoprotein; Membrane; Repeat;
 29 Alternative splicing; GPI anchor; Signal; Sush1; Polymorphism;
 30 Blood group antigen.
 31 SIGNAL: 1 44
 32 CHAIN: 45 453
 33 PROPEP: 454 481
 34 DOMAIN: 45 95
 35 SUSHI 1.
 36 SUSHI 2.
 37 SUSHI 3.
 38 SUSHI 4.
 39 SUSHI 4.
 40 SER/THR-RICH.
 41 COMPLEMENT DEACT ACCELERATING FACTOR.
 42 REMOVED IN MATURE FORM.
 43
 44 N-LINKED (GLCNAC...) (POTENTIAL).
 45 GPI ANCHOR.
 46 HIGHLY SELECTIVE INHIBIT -> SRPVTQAKMKWCURSL.
 47 CSRTQPKRSPHSLSSWYRAHVFHVFADANHCIA
 48 DLAKELRRKRYTOYRIKFIYS (IN ISOFORM 1).
 49 R -> L (IN TC(B) ANTIGEN).
 50 /FTid-VAR_001997.
 51
 52
 53
 54
 55
 56
 57
 58
 59
 60
 61
 62
 63
 64
 65
 66
 67
 68
 69
 70
 71
 72
 73
 74
 75
 76
 77
 78
 79
 80
 81
 82
 83
 84
 85
 86
 87
 88
 89
 90
 91
 92
 93
 94
 95
 96
 97
 98
 99
 100
 101
 102
 103
 104
 105
 106
 107
 108
 109
 110
 111
 112
 113
 114
 115
 116
 117
 118
 119
 120
 121
 122
 123
 124
 125
 126
 127
 128
 129
 130
 131
 132
 133
 134
 135
 136
 137
 138
 139
 140
 141
 142
 143
 144
 145
 146
 147
 148
 149
 150
 151
 152
 153
 154
 155
 156
 157
 158
 159
 160
 161
 162
 163
 164
 165
 166
 167
 168
 169
 170
 171
 172
 173
 174
 175
 176
 177
 178
 179
 180
 181
 182
 183
 184
 185
 186
 187
 188
 189
 190
 191
 192
 193
 194
 195
 196
 197
 198
 199
 200
 201
 202
 203
 204
 205
 206
 207
 208
 209
 210
 211
 212
 213
 214
 215
 216
 217
 218
 219
 220
 221
 222
 223
 224
 225
 226
 227
 228
 229
 230
 231
 232
 233
 234
 235
 236
 237
 238
 239
 240
 241
 242
 243
 244
 245
 246
 247
 248
 249
 250
 251
 252
 253
 254
 255
 256
 257
 258
 259
 260
 261
 262
 263
 264
 265
 266
 267
 268
 269
 270
 271
 272
 273
 274
 275
 276
 277
 278
 279
 280
 281
 282
 283
 284
 285
 286
 287
 288
 289
 290
 291
 292
 293
 294
 295
 296
 297
 298
 299
 300
 301
 302
 303
 304
 305
 306
 307
 308
 309
 310
 311
 312
 313
 314
 315
 316
 317
 318
 319
 320
 321
 322
 323
 324
 325
 326
 327
 328
 329
 330
 331
 332
 333
 334
 335
 336
 337
 338
 339
 340
 341
 342
 343
 344
 345
 346
 347
 348
 349
 350
 351
 352
 353
 354
 355
 356
 357
 358
 359
 360
 361
 362
 363
 364
 365
 366
 367
 368
 369
 370
 371
 372
 373
 374
 375
 376
 377
 378
 379
 380
 381
 382
 383
 384
 385
 386
 387
 388
 389
 390
 391
 392
 393
 394
 395
 396
 397
 398
 399
 400
 401
 402
 403
 404
 405
 406
 407
 408
 409
 410
 411
 412
 413
 414
 415
 416
 417
 418
 419
 420
 421
 422
 423
 424
 425
 426
 427
 428
 429
 430
 431
 432
 433
 434
 435
 436
 437
 438
 439
 440
 441
 442
 443
 444
 445
 446
 447
 448
 449
 450
 451
 452
 453
 454
 455
 456
 457
 458
 459
 460
 461
 462
 463
 464
 465
 466
 467
 468
 469
 470
 471
 472
 473
 474
 475
 476
 477
 478
 479
 480
 481
 482
 483
 484
 485
 486
 487
 488
 489
 490
 491
 492
 493
 494
 495
 496
 497
 498
 499
 500
 501
 502
 503
 504
 505
 506
 507
 508
 509
 510
 511
 512
 513
 514
 515
 516
 517
 518
 519
 520
 521
 522
 523
 524
 525
 526
 527
 528
 529
 530
 531
 532
 533
 534
 535
 536
 537
 538
 539
 540
 541
 542
 543
 544
 545
 546
 547
 548
 549
 550
 551
 552
 553
 554
 555
 556
 557
 558
 559
 560
 561
 562
 563
 564
 565
 566
 567
 568
 569
 570
 571
 572
 573
 574
 575
 576
 577
 578
 579
 580
 581
 582
 583
 584
 585
 586
 587
 588
 589
 590
 591
 592
 593
 594
 595
 596
 597
 598
 599
 600
 601
 602
 603
 604
 605
 606
 607
 608
 609
 610
 611
 612
 613
 614
 615
 616
 617
 618
 619
 620
 621
 622
 623
 624
 625
 626
 627
 628
 629
 630
 631
 632
 633
 634
 635
 636
 637
 638
 639
 640
 641
 642
 643
 644
 645
 646
 647
 648
 649
 650
 651
 652
 653
 654
 655
 656
 657
 658
 659
 660
 661
 662
 663
 664
 665
 666
 667
 668
 669
 670
 671
 672
 673
 674
 675
 676
 677
 678
 679
 680
 681
 682
 683
 684
 685
 686
 687
 688
 689
 690
 691
 692
 693
 694
 695
 696
 697
 698
 699
 700
 701
 702
 703
 704
 705
 706
 707
 708
 709
 710
 711
 712
 713
 714
 715
 716
 717
 718
 719
 720
 721
 722
 723
 724
 725
 726
 727
 728
 729
 730
 731
 732
 733
 734
 735
 736
 737
 738
 739
 740
 741
 742
 743
 744
 745
 746
 747
 748
 749
 750
 751
 752
 753
 754
 755
 756
 757
 758
 759
 760
 761
 762
 763
 764
 765
 766
 767
 768
 769
 770
 771
 772
 773
 774
 775
 776
 777
 778
 779
 780
 781
 782
 783
 784
 785
 786
 787
 788
 789
 790
 791
 792
 793
 794
 795
 796
 797
 798
 799
 800
 801
 802
 803
 804
 805
 806
 807
 808
 809
 810
 811
 812
 813
 814
 815
 816
 817
 818
 819
 820
 821
 822
 823
 824
 825
 826
 827
 828
 829
 830
 831
 832
 833
 834
 835
 836
 837
 838
 839
 840
 841
 842
 843
 844
 845
 846
 847
 848
 849
 850
 851
 852
 853
 854
 855
 856
 857
 858
 859
 860
 861
 862
 863
 864
 865
 866
 867
 868
 869
 870
 871
 872
 873
 874
 875
 876
 877
 878
 879
 880
 881
 882
 883
 884
 885
 886
 887
 888
 889
 890
 891
 892
 893
 894
 895
 896
 897
 898
 899
 900
 901
 902
 903
 904
 905
 906
 907
 908
 909
 910
 911
 912
 913
 914
 915
 916
 917
 918
 919
 920
 921
 922
 923
 924
 925
 926
 927
 928
 929
 930
 931
 932
 933
 934
 935
 936
 937
 938
 939
 940
 941
 942
 943
 944
 945
 946
 947
 948
 949
 950
 951
 952
 953
 954
 955
 956
 957
 958
 959
 960
 961
 962
 963
 964
 965
 966
 967
 968
 969
 970
 971
 972
 973
 974
 975
 976
 977
 978
 979
 980
 981
 982
 983
 984
 985
 986
 987
 988
 989
 990
 991
 992
 993
 994
 995
 996
 997
 998
 999
 1000

FT VARIANT 52 52 R -> P (IN TC(C) ANTIGEN).
 FT VARIANT 82 82 L -> R (N WES(A) ANTIGEN).
 FT VARIANT 199 199 /FTid-VAR_001997.
 FT VARIANT 227 227 S -> L (N DR(A) ANTIGEN).
 FT VARIANT 227 227 A -> P (N CR(A) ANTIGEN).
 FT CONFLICT 80 80 T -> I (N REF 1, 2 AND 4).
 FT CONFLICT 85 85 S -> M (N REF 3).
 FT CONFLICT 187 187 S -> I (N REF 4).
 FT CONFLICT 297 297 Q -> H (IN REF 4).
 SQ SEQUENCE 361 AA: 41388 MW: 2914832B643665E (CR64).
 Query Match: 21.18; Score: 103.5; DB: 1; Length: 481;
 Best Local Similarity: 33.38; Pred. No: 8.8; 09;
 Matches: 45 Conservative 17; Mismatches 50; Indels 24; Gaps 9;
 QY 3 SCGS-PEITLGRISYSTP--IAVTVIRYS--SGTEFLIGKSLNITIKQVIRT--WD 57
 DB 162 SCNPGEITNCOI---DVPGLLPCATLPSNIGYKLPQSTSSPEL- --ISCSWQWS 214
 QY 58 KIAPACIEFNKYSSCPPIVPGCYKIRCSHP-YRHGSVIAFAKINSMGNKSNVQDAN 116
 DB 215 DLPPE---PELYCPAPPPQIDNGLIQ-EEIYGYRQSVTYA-NKGFPMIEHSITIVN 270
 QY 117 N---MW-GPTPLPQC 127
 DB 271 NDECFKSGHP--PQC 284
 RESULT 17
 CQPH_HSVSA
 ID CQPH_HSVSA STANDARD; PR: 360 AA.
 AC Q01016;
 DT 01-APR-1994 (Rel. 25, Created)
 DT 01-APR-1994 (Rel. 25, Last sequence update)
 DT 15-JUL-1998 (Rel. 36, Last annotation update)
 DE Complement control protein homolog precursor (CQPH).
 GS 4 OR CQPH.
 OS Herpesvirus saimiri (strain 11).
 OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
 OC Gammaherpesvirinae; Rhadinovirus
 OX NCBI_TaxID:10383;
 RN 111
 RP SEQUENCE FROM N.A.
 PX MEDLINE-92-33688; PubMed-1321287
 IA Albrecht J.-C., Nicholas J., Biller D., Cameron K.R., Briesinger R.,
 FA Newman C., Wittmann S., Craxton M.A., Coleman H., Fleckenstein R.,
 RA Honess R.W.
 R* "Primary structure of the herpesvirus saimiri genome."
 RL J. Virol. 66:5047-5058(1992)
 RN 121
 RP SIMILARITY TO CCP.
 RX MEDLINE-92260674; PubMed-1316492
 RA Albrecht J.-C., Fleckenstein B.,
 RT herpesvirus saimiri.";
 RL J. Virol. 66:3937-3940(1992).
 CC -I- ALTERNATIVE PRODUCTS: A MEMBRANE-BOUND FORM AND A SECRETED FORM
 CC ARE PRODUCED BY ALTERNATIVE SPLICING OF THE SAME GENE.
 CC -I- SIMILARITY: BELONGS TO THE SUPERFAMILY OF THE REGULATORS OF
 CC COMPLEMENT ACTIVATION (RCA).
 CC -I- SIMILARITY: CONTAINS 3 SUSHI (SCR) DOMAINS.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announcement/>
 CC or send an email to license@sib-sib.ch).
 CC

this SWISS prot entry is copyrighted. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation at the European Bioinformatics Institute. There are no restrictions on its use by non profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announcement/> or send an email to license@isb-sib.ch).

EMBL: L11244; AAA0615.1;
EMBL: L11244; AAA0616.1;
EMBL: M29964; AAB09520.1;
PIR: A44877; A48877.
PIR: A47107; A47107.
ESSE: P10998; IVD.
MIM: 120831;
InterPro: IPRO000436; Sushi scr ctf.
Pfam: PF00084; sushi_3.
SMART: SMO0042; CTP; 4.
Complement pathway; Plasma; Glycoprotein; Repeat; Sushi; Signal;
Polymorphism.
KW SIGNAL. 1 17
FT CHAIN 18 252
FT D-MAIN 22 77
FT D-MAIN 60 135
FT D-MAIN 138 192
FT DISULFID 24 64
FT DISULFID 49 76
FT DISULFID 81 121
FT DISULFID 107 134
FT DISULFID 139 179
FT DISULFID 165 191
FT DISULFID 202
FT DISULFID 216
FT DISULFID 216
FT CARBOHYD 64 64
FT CARBOHYD 71 71
FT CARBOHYD 98 98
FT CARBOHYD 117 117
FT CARBOHYD 154 154
FT VARIANT 198 198
SEQUENCE 252 AA: 28457 MW: 697064367C5E2E7 CRC6A,
Query Match 18 RW: Score 141; DB 1; Length 252;
Best Local Similarity 26.0%; Pred No. 3,4e-07;
Matches 43; Conservative 21; Mismatches 59; Indels 14; Gaps 6;
27 4 CSNPTPLNCRISYSYSLTAVGVV KYSSGIFRLGRKSNDDLLIKPVGLWDAAPK 62
28 ILEI - - - - - I - - - - - I : : : : :
29 CPELPVN - STFAKEVERQITIGTYVCIRGYHLVGKKTLFC --- NASKENWTTE 75
30 CEYNKYSSCDPEIVPRCKRYRGSITFYRHGDVSATACNTFSMNCKNSVMCOANNWCPT 122
31 ILEI : : : : : I - - - - - I : : : : :
32 RICHCHDVVLWG-LFSSSGLVNVSDUKITFMCHDYILKGSKNSQCLEDHTWAPP 130
33 KLETCS 129
34 FPKCS 136
RESULT 26
FEAB FIG STANDARD; PET: 151 AA.
AC Q04710;
15 DEC 1998 (Rel. 37, created)
15 DEC 1998 (Rel. 37, last sequence update)
16 OCT 2001 (Rel. 40, last annotation update)
HE Complement factor B (E: 3.4.21.47) (3/CS convertase) (Properdin
factor B) (fragment).
GN Sus scrofa (pig)

[illegible]

RP VARIANTS C2D, PHE-209 AND ARG-464.
RX MEDLINE-96215049; PubMed-8621452;
RA Wotzel R.A., Kullis J., Iwakura M., Kiopele P., Akama H.,
RA Johnson C.A., Jensen P., Goffon H.P.;
RT "Type 11 human complement C2 deficiency. Allele-specific amino acid
RT substitutions (Ser186 -> Phe; Gly444 -> Arg) cause impaired C2
RT secretion.";
RL J. Biol. Chem. 271:5824-5831(1996).
RN [9]
RV VARIANT C2D, IYR-131.
RX MEDLINE-98344005; PubMed-9670940;
RA Zhu Z.B., Atkinson T.P., Watanabe T.;
RA "A novel type II complement C2 deficiency allele in an African
RT American family.";
RL J. Immunol. 161:578-584(1998).
CC 1- FUNCTION: COMPONENT C2 WHICH IS PART OF THE CLASSICAL PATHWAY OF
CC THE COMPLEMENT SYSTEM IS CLEAVED BY ACTIVATED FACTOR C3 INTO TWO
CC FRAGMENTS: C2B AND C2A. C2A, A SERINE PROTEASE, THEN COMBINES WITH
CC COMPLEMENT FACTOR 4B TO GENERATE THE C3 CONVERTASE
CC 1- CATALYTIC ACTIVITY: CLEAVES C3 IN THE ALPHA-CHAIN TO YIELD C3A AND
CC C3B. CLEAVES C5 IN THE ALPHA-CHAIN TO YIELD C5A AND C5B. BOTH
CC CLEAVAGES TAKE PLACE AT THE C-TERMINAL OF AN ARGININE RESIDUE
CC 1- DISEASE: DEFECTS IN C2 ARE THE CAUSE OF C2 DEFICIENCY (C2D). THIS
CC IS AN AUTOIMMUNE RECURRING DISEASE. DEFICIENT INDIVIDUALS HAVE AN
CC INCREASED INCIDENCE OF SLE AND SLE-LIKE SYNDROMES.
CC GLOMERULONEPHRITIS, VASCULITIS AND PYOGENIC INFECTIONS. TYPE 1 C2D
CC IS CHARACTERIZED BY COMPLETE LOSS OF THE PROTEIN WHILE TYPE 11 C2D
CC IS CHARACTERIZED BY A SELECTIVE BLOCK IN C2 SYNTHESIS.
CC 1- MISCELLANEOUS: C2 IS A MAJOR HISTOCOMPATIBILITY COMPLEX CLASS-III
CC PROTEIN.
CC 1- SIMILARITY: WITH COMPLEMENT FACTOR B.
CC 1- SIMILARITY: CONTAINS 4 SUSHI (SCR) DOMAINS.
CC 1- SIMILARITY: CONTAINS 1 VWFA DOMAIN.
CC 1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1; ALSO KNOWN AS THE
CC TRYPSIN FAMILY.
CC
CC This SWISS PROT entry is copyright 1998, is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by not-for-profit institutions as long as its content is in no way
CC modified and this statement is not removed, usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or
CC or send an email to license@isb-sib.ch).
DR EMBL: M15082; AAA59624.1;
DR EMBL: L09708; AAB97607.1;
DR EMBL: L09706; AAB97607.1; JOINED.
DR EMBL: L09707; AAB97607.1; JOINED.
DR EMBL: AF019416; AAB67975.1;
DR EMBL: X04481; CAA28169.1;
DR EMBL: A25290; C2H0.
DR EMBL: A05289; A05289.
DR DSSP: P00744; 1VW1.
DR MEROPS: S01.194;
DR MIM: 217000;
DR InterPro: IPR001414; Chymotrypsin.
DR InterPro: IPR000436; Sushi Scr Dep.
DR InterPro: IPR001254; Trypsin.
DR InterPro: IPR002045; VWFA.
DR Pfam: Pfam0084; sushi.
DR Pfam: Pfam0089; trypsin; 2.
DR Pfam: Pfam0092; vwa; 1.
DR PRINTS: PR00722; CHYMOTRYPSIN.
DR SMART: SM00042; C2P; 3.
DR SMART: SM00020; TRYP-SER; 1.
DR SMART: SM00327; VWA; 1
DR PROSITE: PS02240; TRYP-SIN-DEM; 1.
DR PROSITE: PS00134; TRYP-SIN-HIS; 1.
DR PROSITE: PS00135; TRYP-SIN-SER; 1.
DR PROSITE: PS02244; VWFA; 1.
KW Complement pathway; Plasma: Glycoprotein; Hydrolase; Serine protease
KW Signal; Repeat; Sushi; Disease mutation; Polymorphism.
FT SIGNAL 1 20

FT CHAIN 21 752
FT CHAIN 21 343
FT CHAIN 244 752
FT D-MAIN 23 85
FT D-MAIN 88 145
FT D-MAIN 50 205
FT D-MAIN 254 452
FT D-MAIN 466 752
FT ACT-SITE 507 507
FT ACT-SITE 561 561
FT ACT-SITE 579 679
FT DISULFID 24 64
FT DISULFID 51 84
FT DISULFID 89 131
FT DISULFID 117 144
FT DISULFID 151 191
FT DISULFID 177 204
FT CARBOHYD 23 29
FT CARBOHYD 112 112
FT CARBOHYD 290 290
FT CARBOHYD 333 333
FT CARBOHYD 467 467
FT CARBOHYD 471 471
FT CARBOHYD 621 621
FT CARBOHYD 651 651
FT VARIANT 131 131
FT VARIANT 209 209
FT VARIANT 464 464
FT VARIANT 533 533
FT SEQUENCE 752 AA; 83267 MW; 5A96A18E700CF444 CRC64;
Query Match 18.6%; Score 140; IB 1; Length 752;
Best Local Similarity 28.5%; Prec No. 1.4e-06;
Matches 37; Conservative 16; Mismatches 61; Indels 16; Gaps 6.
QY 2 ISGSHPPPLNG---RISYYSIPIAVG VIFYSYSCSIFRLGKSKSLCIIKKKVRILWD 57
DB 87 VRCPAPVFENGIYTPRLQSY-----PVCINV; FECEGFIILGSPVPGG-----KPNQMD 148
QY 58 KPAFKCEYNNKYSSCPETVPGYKIRGSTYRHGDSVTFACKTNFNMGNKSNVWVQANN 117
DB 139 GETAVCA--NCAGCHQNPICISUG-AVRGSPFGHGWVPPVHCSSNIVLQSSSRERFGN; 195
QY 118 MKGFTGLPTC 127
DB 196 VWSGTE-PLC 204
RESULT 24
APOL_MOUSE
ID APOL_MOUSE STANDARD: PPI: 345 AA.
AC Q01339;
DI 31-APR-1993 (Rel. 25, Created)
DI 31-APR-1993 (Rel. 25, Last sequence update)
DI 16-OCT-2001 (Rel. 40, Last annotation update)
DE Beta-2-glycoprotein I precursor (APolipoprotein I) (Apo. II) (B2GPI)
DE (Beta(2)GPI) (Activated protein C-binding protein) (APC inhibitor).
GN APOII OR B2GPI.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Insecta; Rodentia; Sciurgnathia; Muridae; Mus.
OX NCBI_TaxID:10090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE-92372000; PubMed-1334387;
RA Nonaka M., Matsuda Y., Shirotshi T., Moriwa K., Natsume Sakai S.;
RT "Molecular cloning of mouse beta 2-glycoprotein I and mapping of the
RT gene to chromosome 11.";
RL Genomics 14:1082-1087(1992).

Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo; NCBI TaxID: 9606;

SEQUENCE FROM N.A. (ALLEN S); FA AND FB); MEDLINE 91065702; PubMed 2249879;

DAVIDSON J, ALLEN M, CLARK A; Molecular characterization of human complement factor B subtypes; Immunogenetics 32:309-312(1990);

SEQUENCE FROM N.A. (ALLEN S);

TISSUE-LIVER;

MEDLINE 9423735; PubMed 8319623;

Mejia J.E., Jahn T., de la Salla H., Hauptmann G.; Human factor B: Complete cDNA sequence of the F4S allele; Hum. Immunol. 49:49-53(1994);

SEQUENCE FROM N.A. (ALLEN S);

TISSUE-LIVER;

MEDLINE 94041499; PubMed 8225486;

Schwartz W., Lucifora H., Sokolowski L., Esteller C., Weiss E.H.; Human complement factor B: functional properties of a recombinant enzyme of the alternative activation pathway convertase; J. Immunobiology 188:223-232(1993);

SEQUENCE FROM N.A. (ALLEN S);

MEDLINE 94067477; PubMed 8247029;

Horiochi T., Kim S., Matsuno M., Watanabe T., Fujita S.; Human complement factor B: cDNA cloning, nucleotide sequencing, phenotypic conversion by site directed mutagenesis and expression; Mol. Immunol. 30:1587-1592(1994);

SEQUENCE FROM N.A. (ALLEN S);

Rowen L., Bankers C., Buskin D., Faust J., Lorent, C., Alcantara M.E., Bonta A., Swartzell S., Smith T.M., Spies T., Boed L.; Sequences determination of 300 kilobases of the human class III MHC locus; Submitted (SEP-1997) to the EMBL/Genbank/DBJ databases;

SEQUENCE FROM N.A. (ALLEN S);

Strausberg R.; Submitted (MAR-2001) to the EMBL/Genbank/DBJ databases;

SEQUENCE FROM N.A. (ALLEN S);

Bohnen T., Kuehn J., Fontana K.E., Saarinen-Pihkala B., Lakk M.; Expression and alternative splicing of human factor B gene in leukemic mononuclear cells; Submitted (FEB-2001) to the EMBL/Genbank/DBJ databases;

SEQUENCE FROM N.A. (ALLEN S);

MEDLINE 84161997; PubMed 6346754;

Mele J.E., Anderson J.K., Davidson E.A., Woods D.E.; Complete primary structure for the zymogen of human complement factor B; J. Biol. Chem. 259:3407-3412(1984);

SEQUENCE FROM N.A. (ALLEN S);

MEDLINE 84204002; PubMed 6342610;

Christie J., Garton J.; Amino acid sequence of the Bb fragment from complement factor B;

100 Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo; NCBI TaxID: 9606;

101 SEQUENCE FROM N.A. (ALLEN S); FA AND FB); MEDLINE 91065702; PubMed 2249879;

102 DAVIDSON J, ALLEN M, CLARK A; Molecular characterization of human complement factor B subtypes; Immunogenetics 32:309-312(1990);

103 SEQUENCE FROM N.A. (ALLEN S);

104 TISSUE-LIVER;

105 MEDLINE 9423735; PubMed 8319623;

106 Mejia J.E., Jahn T., de la Salla H., Hauptmann G.; Human factor B: Complete cDNA sequence of the F4S allele; Hum. Immunol. 49:49-53(1994);

107 SEQUENCE FROM N.A. (ALLEN S);

108 TISSUE-LIVER;

109 MEDLINE 94041499; PubMed 8225486;

110 Schwartz W., Lucifora H., Sokolowski L., Esteller C., Weiss E.H.; Human complement factor B: functional properties of a recombinant enzyme of the alternative activation pathway convertase; J. Immunobiology 188:223-232(1993);

111 SEQUENCE FROM N.A. (ALLEN S);

112 MEDLINE 94067477; PubMed 8247029;

113 Horiochi T., Kim S., Matsuno M., Watanabe T., Fujita S.; Human complement factor B: cDNA cloning, nucleotide sequencing, phenotypic conversion by site directed mutagenesis and expression; Mol. Immunol. 30:1587-1592(1994);

114 SEQUENCE FROM N.A. (ALLEN S);

115 Rowen L., Bankers C., Buskin D., Faust J., Lorent, C., Alcantara M.E., Bonta A., Swartzell S., Smith T.M., Spies T., Boed L.; Sequences determination of 300 kilobases of the human class III MHC locus; Submitted (SEP-1997) to the EMBL/Genbank/DBJ databases;

116 SEQUENCE FROM N.A. (ALLEN S);

117 Strausberg R.; Submitted (MAR-2001) to the EMBL/Genbank/DBJ databases;

118 SEQUENCE FROM N.A. (ALLEN S);

119 Bohnen T., Kuehn J., Fontana K.E., Saarinen-Pihkala B., Lakk M.; Expression and alternative splicing of human factor B gene in leukemic mononuclear cells; Submitted (FEB-2001) to the EMBL/Genbank/DBJ databases;

120 SEQUENCE FROM N.A. (ALLEN S);

121 MEDLINE 84161997; PubMed 6346754;

122 Mele J.E., Anderson J.K., Davidson E.A., Woods D.E.; Complete primary structure for the zymogen of human complement factor B; J. Biol. Chem. 259:3407-3412(1984);

123 SEQUENCE FROM N.A. (ALLEN S);

124 MEDLINE 84204002; PubMed 6342610;

125 Christie J., Garton J.; Amino acid sequence of the Bb fragment from complement factor B;


```

FT DISULFID 163 163 INTERCHAIN (BY SIMILARITY)
FT CONFLICT 44 44 1 - V (IN REF. 1; AA SEQUENCE)
FT CONFLICT 46 46 1 - A (IN REF. 1; AA SEQUENCE)
SQ SEQUENCE 202 AA; 22724 MW; EABC951B06A0D80C CRC64;

Query Match 17.4%; Score 131; DB 1; Length 202;
Best Local Similarity 25.2%; Pred. N. 2.7e-06;
Matches 34; Conservative 24; Mismatches 56; Indels 22; Caps 6;

QY 4 GSPPPILNGRLSYSTPIAGTV-----IRYSSTGTFRLIGELKSLICITKDKVDGTWD 57
DB 30 CNPFWVARG-----HTTQI-IGLFOMKKDEVVYKDCGYTLVGHDRISCKS-----SRWS 79
QY 58 KPAPEYENKYSSTPEIVAGKXKRGSTPYRHGDSVTFACKTNFNMCKNSKSWCCANN 117
DB 80 PAAPQK-----ALGPKPIQ6RLSVIQDEYIESENVIVGCSGSGYLVGPKLITCTEDG 134
QY 118 MWGPTRLTNSVFP 132
DB 135 TWHP RVKQWEXP 148

RESULT 46
ID CFAB ROVIN STANDARD; PRI: 685 AA;
AC Q28085;
DT 01-MAR-2002 (Rel. 41, Created)
DT 01-MAR-2002 (Rel. 41, Last sequence update)
DI 01-MAR-2002 (Rel. 41, Last annotation update)
DE Complement factor H (H factor 1) (Fragments);
GN HFI;
OS Bos taurus (Bovine);
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Bovinae; Bos;
OX NCBI_TaxID=9913;
RN 111
RP SEQUENCE OF 17 685 FROM N.A. AND SEQUENCE OF 1-16
RT TISSUE=Liver;
RX MEDLINE=96202005; PubMed=8615824;
RA Soares G.J., Day A.J., Sim R.B.;
RI "Prediction from sequence comparisons of residues of factor H involved
RI in the interaction with complement component C3b.";
RL Biochem. J. 315:523-531(1996);
CC 1- FUNCTION: Factor H functions as a cofactor in the inactivation of
CC C3b by factor I and also increases the rate of dissociation of the
CC C3bB complex (C3 convertase) and the (C3b)BBB complex (C5
CC convertase) in the alternative complement pathway (By similarity).
CC 1- SIMILARITY: CONTAINS AT LEAST 13 SUSHI (SCK) DOMAINS.
CC
CC This SWISS Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed, usurped, by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
DB EMBL: X98697; CAA67257.1;
DB RSSP: P10998; IYVD;
DB InterPro: IPR000436; Sushi_SCR_CCP;
DB Pfam: PF00084; sushi; 11;
DB SMART: SM00042; CCP; 11;
KW Complement alternate pathway; Plasma; Repeat; Sushi;
FT NON_TER 1 1
FT NON_TER 16 17
FT DOMAIN 17 67 SUSHI 2;
FT DOMAIN 68 70 SUSHI 3;
FT DOMAIN 71 188 SUSHI 4;
FT DOMAIN 191 246 SUSHI 5;
FT DOMAIN 249 409 SUSHI 6;
FT DOMAIN 411 466 SUSHI 7;
FT DOMAIN 470 429 SUSHI 8;

```

```

FT DOMAIN 431 488 SUSHI 9;
FT DOMAIN 491 547 SUSHI 10;
FT DOMAIN 553 609 SUSHI 11;
FT DOMAIN 614 668 SUSHI 12;
FT DOMAIN 675 >685 SUSHI 13;
FT DISULFID 39 66 HY SIMILARITY;
FT DISULFID 71 117 HY SIMILARITY;
FT DISULFID 103 130 HY SIMILARITY;
FT DISULFID 135 176 HY SIMILARITY;
FT DISULFID 162 187 HY SIMILARITY;
FT DISULFID 192 234 HY SIMILARITY;
FT DISULFID 219 245 HY SIMILARITY;
FT DISULFID 250 297 HY SIMILARITY;
FT DISULFID 280 308 HY SIMILARITY;
FT DISULFID 312 354 HY SIMILARITY;
FT DISULFID 339 365 HY SIMILARITY;
FT DISULFID 371 417 HY SIMILARITY;
FT DISULFID 400 428 HY SIMILARITY;
FT DISULFID 432 476 HY SIMILARITY;
FT DISULFID 459 487 HY SIMILARITY;
FT DISULFID 492 534 HY SIMILARITY;
FT DISULFID 520 546 HY SIMILARITY;
FT DISULFID 554 597 HY SIMILARITY;
FT DISULFID 583 608 HY SIMILARITY;
FT DISULFID 615 656 HY SIMILARITY;
FT DISULFID 642 667 HY SIMILARITY;
FT NON_TER 685
SQ SEQUENCE 685 AA; 77536 MW; 69FC9_C8D530EB72 CRC64;

Query Match 17.4%; Score 141; DB 1; Length 685;
Best Local Similarity 28.5%; Pred. N. 1e-05;
Matches 39; Conservative 19; Mismatches 49; Indels 40; Caps 8;

QY 3 SNGSPPIING-RISYYSPIAGVGLIVYSSTGTFRLIGKSLICITKDKVDGTWIKPAP 61
DB 553 SCALPPLNGKRLKKEHYAHNEVYVYACPKPKGSHKIQG---VDRP-STALP 606
QY 62 KQYFNKYSSCP-----PPVDSQKPIFNSIPYRGNVIFAKKRIEMRCKGVW 112
DB 607 VC TEERTGELSDLDGDKVPSVP-----PIHDSVSESKREAFIMGIRFII 656
QY 113 QVANNWGHTRLPQVS 129
DB 656 C-ISTPW---TQDPQCTA 669

RESULT 37
ID LEM3 SHEEP STANDARD; PRI: 769 AA;
AC P98109;
DT 01-FEB-1996 (Rel. 33, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE P-selectin precursor (Granule membrane protein 140) (GMP-140) (PAK3DM)
DE (CD62P) (Leukocyte endothelial cell adhesion molecule 4) (LECAM4)
GN SELP;
OS Ovis aries (Sheep);
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Hovidae; Caprinae; Ovis;
OX NCBI_TaxID=9940;
RN 111
RP SEQUENCE FROM N.A.
RT TISSUE=Heart;
RA Burns S.A., Neufeld E.J., Donady J.;
DI Submitted (JUL-1994) to the EMBL/Genbank/DBJ databases
DE 1- FUNCTION: CA(2+)-DEPENDENT RECEPTOR FOR WYELOID CELLS THAT BINDS
DE TO CARBOHYDRATES ON NEUTROPHILS AND MONOCYTES. MEDIATES THE
DE INTERACTION OF ACTIVATED ENDOTHELIAL CELLS OR PLATELETS WITH
DE LEUKOCYTES. THE LIGAND RECOGNIZED IS STAYL LEWIS X.
CC 1- SUBCELLULAR LOCATION: TYPE I membrane protein.
CC 1- SIMILARITY: TO OTHER SELECTINS/LECAMs
CC 1- SIMILARITY: CONTAINS 1 C-TYPE JETIN FAMILY DOMAIN.

```

	604	604	N LINKED (GLYNA ⁺ ...) (P ⁺ ENTIAL)
FT	CARBOHYD	604	N LINKED (GLYNA ⁺ ...) (P ⁺ ENTIAL)
FT	CARBER HIN	604	N LINKED (GLYNA ⁺ ...) (P ⁺ ENTIAL)
FT	CARBOHYD	604	N LINKED (GLYNA ⁺ ...) (P ⁺ ENTIAL)
FT	CARBOHYD	604	N LINKED (GLYNA ⁺ ...) (P ⁺ ENTIAL)
FT	SLIFE	757	N LINKED (GLYNA ⁺ ...) (P ⁺ ENTIAL)
FT	VARIANT	606	ENDONYMUS SIGNAL (P ⁺ ABLE)
FT	VARIANT	679	S...T...
FT	VARIANT	679	T...V...
SQ	SEQUENCE	769 AA	84.17 MW; 244,267 Da (Molecular Weight)

```

BEST LOCAL SIMILARITY 25.187, ID: 60, LEN: 69,
Matches 49, Conservative 18, Mismatches 59, Indels 18, Gaps 7.

QY 1 GIGVSGSPFPLENGRIS--VYV--PIAVVIVIVYVSSGSPFPEKPSKSLVLEIVVLEIVLE 58
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 578 GLRGSLSIAPNCHVNSCHHRLRNSGLNLTTHPGVKKAFILMGESALQC  KPSKOWTA 633
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 59 PAVVTEVTPVPSYSSV--FLAVV--VYVLEIVLEIV  VYVTEVYVAVPPIRPPGRSVWV 114
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 634 VAPTEPAV KSKSLVTEPIV--KMSNPVNTSVSSPSPVTEALLNLSRTAAV 687
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 114 QANNMCHVIRLPIIC 127
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 688 QFNQWS-TTETC 700
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

RESULT 38
LEM2_RAT
ID LEM2_RAT STANDARD; PRI: 649 AA.
AC P98105;
DT 01-FEB-1996 (Rel. 33, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 01-FEB-1996 (Rel. 33, Last annotation update)
E SELECT-IN precursor (Endothelial leukocyte adhesion molecule-1)
DE (EAM-1) (leukocyte endothelial cell adhesion molecule 2) (LECAM2)
DE (CD62E).
GN SELE OR ELAM-1.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus
OC NCBI:TXID:10116;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE Lung;
RC Rosenblum C.H., Auchampach J.A., Anderson D.C., Mancini A.M.;
RC Submitted (May 1993) to the EMBL/GenBank/DBJ databases.
CC 1- FUNCTION: EXPRESSED ON CYTOKINE INDUCED ENDOTHELIAL CELLS AND
CC MEDIATES THEIR BINDING TO LEUKOCYTES. THE LIGAND RECOGNIZED BY
CC EAM-1 IS SIALYL-Lewis X (SLeX) (GLCNAc1-6)GALACTOSYLATED DERIVATIVES OF
CC POLYLACTOSAMINE THAT ARE FOUND AT THE NONRELATING TERMINUS OF
CC GLYCOPOLYMER.
CC 1- SUBCELLULAR LOCATION: TYPE I membrane protein.
CC 1- SIMILARITY: TO OTHER SELECTINS/LEGAMS.
CC 1- SIMILARITY: CONTAINS 1 C TYPE LECTIN FAMILY 1-MAIN
CC 1- SIMILARITY: CONTAINS 1 EGF-LIKE DOMAIN
CC 1- SIMILARITY: CONTAINS 5 SUSHI (SCR) DOMAINS.
CC
CC THIS SWISS-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC at the European Bioinformatics Institute. There are no restrictions on its
CC use by non profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement. For more information, please
CC or send an email to license@isb.sdb.ch).
CC
CC EMBL: L25527; AAA41113.1;
CC USSP: P16681; IKJA.
DR InterPro: IPR000561; EGF-like.
DR InterPro: IPR000436; Sushi_Scr_Ccp.
DR InterPro: IPR001304; Lectin_C.
DR Pfam: PF00008; EGF; 1.
DR Pfam: PF00059; Lectin_C; 1.
DR Pfam: PF00084; sushi; 5.

```


1 CC 1 SIMILARITY: CONTAINS 6 SUSHI (SUS) DOMAINS: BOVINE P LECTIN LACKS
2 THE HUMAN SUSHI 3, 4 AND 7 EQUIVALENTS
3
4 CC This SWISS-Prot entry is copyright. It is produced through a collaboration
5 between the Swiss Institute of Bioinformatics and the EMBL outstation
6 at the European Bioinformatics Institute. There are no restrictions on its
7 use by non-profit institutions as long as its content is in no way
8 modified and this statement is not removed. Usage by and for commercial
9 entities requires a license agreement (See <http://www.ebi.ac.uk/submit/faq/>
10 or send an email to license@ebi.ac.uk).

10 EMBL: U22041; AAA02431; 1
11 RESPI: P16109; LF58
12 InterPro: IPR000561; EGF like.
13 InterPro: IPR002496; Selectin.
14 InterPro: IPR003446; SUSHI SUS EXP.
15 InterPro: IPR001404; Lectin c.
16 Pfam: PR00008; EGF 1
17 Pfam: PR00059; Lectin c. 1
18 Pfam: PR00064; SUSHI 6
19 PRINTS: PR00344; SELECTIN.
20 SMART: SM00042; CYP 6.
21 SMART: SM00044; CLECT 1.
22 SMART: SM00181; EGF 1.
23 PROSITE: PS00022; EGF 1.
24 PROSITE: PS00163; EGF 2.
25 PROSITE: PS00041; C TYPE LECTIN 1.
26 PROSITE: PS00041; C TYPE LECTIN 2.
27 K1 Well adhesion. Transmembrane; glycoprotein; EGF like domain; Lectin;
28 Selectin. Signal; SUSHI repeat.
29
30 FT SIGNAL 1 41 BY SIMILARITY.
31 FT CHAIN 42 645 P-SELECTIN.
32 FT DOMAIN 42 645 EXTRA-CELLULAR (POTENTIAL).
33 FT TRANSMEM 588 611 POTENTIAL.
34 FT DOMAIN 612 646 CYTOPLASMIC (POTENTIAL).
35 FT DOMAIN 58 158 C TYPE LECTIN (SHORT FORM).
36 FT DOMAIN 159 195 EGF LIKE.
37 FT DOMAIN 199 258 SUSHI 1.
38 FT DOMAIN 261 320 SUSHI 2.
39 FT DOMAIN 323 382 SUSHI 3.
40 FT DOMAIN 385 444 SUSHI 4.
41 FT DOMAIN 457 516 SUSHI 5.
42 FT DOMAIN 519 578 SUSHI 6.
43 FT DISULFID 60 136 BY SIMILARITY.
44 FT DISULFID 141 196 BY SIMILARITY.
45 FT DISULFID 163 174 BY SIMILARITY.
46 FT DISULFID 188 194 BY SIMILARITY.
47 FT DISULFID 200 244 BY SIMILARITY.
48 FT DISULFID 240 257 BY SIMILARITY.
49 FT DISULFID 262 306 BY SIMILARITY.
50 FT DISULFID 292 319 BY SIMILARITY.
51 FT DISULFID 324 368 BY SIMILARITY.
52 FT DISULFID 354 381 BY SIMILARITY.
53 FT DISULFID 386 430 BY SIMILARITY.
54 FT DISULFID 416 438 BY SIMILARITY.
55 FT DISULFID 458 502 BY SIMILARITY.
56 FT DISULFID 488 515 BY SIMILARITY.
57 FT DISULFID 520 544 BY SIMILARITY.
58 FT DISULFID 550 577 BY SIMILARITY.
59 FT CARBOHYD 48 38 N LINKED (GLNAC).
60 FT CARBOHYD 53 54 N LINKED (GLNAC).
61 FT CARBOHYD 80 80 N LINKED (GLNAC).
62 FT CARBOHYD 180 180 N LINKED (GLNAC).
63 FT CARBOHYD 212 212 N LINKED (GLNAC).
64 FT CARBOHYD 219 219 N LINKED (GLNAC).
65 FT CARBOHYD 336 336 N LINKED (GLNAC).
66 FT CARBOHYD 481 481 N LINKED (GLNAC).
67 FT CARBOHYD 542 542 N LINKED (GLNAC).
68 FT CARBOHYD 549 549 N LINKED (GLNAC).
69 FT CARBOHYD 563 567 N LINKED (GLNAC).
70 SITE 634 647 ENDOXYTOSIS SIGNAL (PROBABILE).
71 SEQUENCE 646 AA: 71229 MW: 57012.440.2/ALATA CR064;
72

Query Match 16.9% Score 127; DB 1; Length 646;
Best Local Similarity 26.9% Evid. No. 2.4e-06;
Matches 46; Conservative 21; Mismatches 59; Indels 18; Gaps 7;
QY 1 GLESGDPPPLNPRIS VYTFIVAVIVVYSSEIFEDKSKDITFKKVIWIK 56
DB 455 GVQGTLLAKKQTSQHVNFQNTCHPKKAGFTLSDALQK RPSROWIA 510
QY 59 FAKEVYTPYSSQF FIVIVGVYFVSEF YEEKCVILVETNEVDRKSNVW 114
DB 511 AADICRAV KCAKLVIRPIV - MNINWGNFSGYSGISFHPDQALINSEIV 564
QY 114 GANNMGPTRPIC 127
DB 565 QENQWS-TIMPTC 577

RESULT 42

LEW2_MOUSE

ID LEW2_MOUSE STANDARD: PRI: 612 AA

AC Q00690;

DT 01-APR-1994 (Rel. 25, Created)

DT 01-APR-1994 (Rel. 25, Last sequence update)

DT 01-FEB-1996 (Rel. 53, Last annotation update)

DE E-selectin precursor (endothelial leukocyte adhesion molecule 1)

DE (ELAM-1) (Leukocyte-endothelial cell adhesion molecule 2) (LECAM2)

DE (CD62E).

GN SELE OR ELAM-1.

OS Mus musculus (Mouse).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Rodentia; Sciurostomatia; Muridae; Mus;

OX NCBI_TaxID=10090;

RN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE:92283265; PubMed:1375914;

RA Hockett Andre M.; van Halbeek H.; Leshetzer C.; Whelan J.

RA Dolanar J.F.;

RT "Marine endothelial leukocyte adhesion molecule 1 is a close

RT structural and functional homologue of the human protein."

RL Eur. J. Biochem. 206:401-411(1992).

RN [2]

RP SEQUENCE FROM N.A.

RX MEDLINE:92440571; PubMed:1378846;

RA Weller A.; Isenmann S.; Vestweber D.

RT "Cloning of the mouse endothelial selectins. Expression of both E

RT and P-selectin is inducible by tumor necrosis factor alpha."

RL J. Biol. Chem. 267:15176-15183(1992).

CC 1- FUNCTION: EXPRESSED ON CYTOKINE INDUCED ENDOTHELIAL CELLS AND

CC MEDIATES THEIR BINDING TO LEUKOCYTES. THE LIGAND RECOGNIZED BY

CC ELAM-1 IS STALYL LEWIS X (ALPHA(1-3)FUCOSYLATED DERIVATIVES OF

CC POLYLACTOSAMINE THAT ARE FOUND AT THE N-TERMINUS TERMINI OF

CC GLYCOPROTEINS).

CC 1- SUBCELLULAR LOCATION: Type I membrane protein.

CC 1- SIMILARITY: TO OTHER SELECTINS/LEWIS.

CC 1- SIMILARITY: CONTAINS 1 C-TYPE LECTIN FAMILY DOMAIN.

CC 1- SIMILARITY: CONTAINS 1 EGF LIKE DOMAIN.

CC 1- SIMILARITY: CONTAINS 6 SUSHI (SUS) DOMAINS.

CC This SWISS-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC at the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.ebi.ac.uk/submit/faq/>
CC or send an email to license@ebi.ac.uk).

DR EMBL: M80778; AAA37547.1; -

DR EMBL: M87862; AAA37577.1; ALT_INIT.

DR ISSP: P16581; IKA.

DR MGI: MGI:98278; Sele.

DR InterPro: IPR000561; EGF like.

DR InterPro: PR002496; Selectin.

DR

[illegible]

DE SMART: SM00044; CLEFT: 1;
 DE SMART: SM00181; EGF: 2;
 DE PROSITE: PS00022; EGF: 1;
 DE PROSITE: PS01186; EGF: 2;
 DE PROSITE: PS00615; C-TYPE_LECTIN_1; 1;
 DE PROSITE: PS50041; C-TYPE_LECTIN_2; 1;
 KW Cell adhesion; Transmembrane; Glycoprotein; EGF-like domain; Lectin;
 KW Selectin; Signal; Sushi; Repeat;
 FT SIGNAL 1 22
 FT CHAIN 23 611
 FT DOMAIN 23 557
 FT TRANSMEM 580 579
 FT DOMAIN 580 611
 FT DOMAIN 39 139
 FT DOMAIN 140 176
 FT DOMAIN 180 239
 FT DOMAIN 242 301
 FT DOMAIN 304 464
 FT DOMAIN 467 427
 FT DOMAIN 430 490
 FT DOMAIN 493 549
 FT DISULFID 41 139
 FT DISULFID 112 141
 FT DISULFID 144 155
 FT DISULFID 149 164
 FT DISULFID 166 176
 FT DISULFID 181 225
 FT DISULFID 211 238
 FT DISULFID 243 287
 FT DISULFID 273 300
 FT DISULFID 305 350
 FT DISULFID 336 363
 FT DISULFID 368 413
 FT DISULFID 399 426
 FT DISULFID 431 476
 FT DISULFID 462 489
 FT DISULFID 494 535
 FT DISULFID 521 548
 FT CARBOHYD 26 26
 FT CARBOHYD 161 161
 FT CARBOHYD 204 204
 FT CARBOHYD 266 266
 FT CARBOHYD 413 413
 FT CARBOHYD 433 433
 FT CARBOHYD 528 528
 FT CARBOHYD 528 528
 SQ SEQUENCE 611 AA; 35DA9E3DF225E4F6 CRC64;
 Query Match 16.2%; Score 122; DB 1; Length 611;
 Best Local Similarity 25.5%; Pred. No. 7.2e 05;
 Matches 6; Conservative 24; Mismatches 55; Indels 26; Gaps 7;
 QY 2 TSAGSPPTLNGFTRSYSTD---LACIVIPYSCSTFFLQKSLGCTTKDKVICTWDR 58
 ID 303 VTGALGHQNGSVSGSHSPAEFSVRSSCNFTCNQEGFLMGSPAICTAQ-----GQMSQ 358
 QY 59 PAIKCFYFNKYSQ--PE--PIVPGYKIKNGSTPYRIGDSVTFACKTNFSMNGK 109
 ID 459 QVFWC--KASQKALSSPERGYNSLPFG----ASQSPQSGSCFFCFCKGFVLKSK 409
 QY 110 SYWQANNMWGTRLPFTVSV 130
 ID 410 TLQGLTGKW DSEETLEAV 429
 RESULT 45
 CFAB MOUSE
 AC 104186;
 DT 20-MAR-1987 (Ref. 04, Treated)
 DT 01-AUG-1997 (Ref. 19, Last sequence update)
 DT 16-DEC-2001 (Ref. 40, Last annotation update)
 DE Complement factor B precursor (EC 3.4.21.27) (c3/c5 convertase).
 GN BF OR B2 BF.

OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurostathi Muridae; Murinae; Mus
 OC NCBI_taxonomy:0090;
 RN [1]
 RN SEQUENCE FROM N.A.
 RX MEDLINE:9,05430; PubMed:2229060;
 RA Ishikawa N., Nonaka M., Weisel R.A., Colter H.R.,
 RT "Murine complement C2 and factor B polymorphic and cDNA cloning reveals
 RT different mechanisms for multiple transcripts of C2 and B2".
 RJ J. Biol. Chem. 265:19040-19046(1990).
 RN [2]
 RN SEQUENCE FROM N.A.
 RA Rowen L., Qin S., Lasky S.R., Loretz C., Dars M., Mahalingam G.,
 RA Hood L.,
 RJ Submitted (Feb-1998) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RN SEQUENCE FROM N.A.
 RA STRAIN-129;
 RA Rowen L., Madan A., Qin S., Hall C., Dahl T., James K., Dickhoff K.,
 RA Schaffer L., Pateliff A., Abbasi N., Loretz C., Lasky S., Hood L.,
 RJ "Sequence of the mouse MHC class III region".
 RJ Submitted (NOV-1998) to the EMBL/GenBank/DBJ databases.
 RN [4]
 RN SEQUENCE OF 285-761 FROM N.A.
 RX MEDLINE-84061931; PubMed-6689022;
 RA Sackstein R., Colten H.R., Woods J.F.,
 RT "Phylogenetic conservation of a class III major histocompatibility
 RT complex antigen, factor B, isolation and nucleotide sequencing of
 RJ mouse factor B cDNA clones.";
 RJ J. Biol. Chem. 258:14693-14697(1983)
 CC -1- FUNCTION: FACTOR B WHICH IS PART OF THE ALTERNATE PATHWAY OF THE
 CC COMPLEMENT SYSTEM IS CLEAVED BY FACTOR D INTO 2 FRAGMENTS: BA AND
 CC BB. BB, A SERINE-PROTEASE, THEN COMBINS WITH COMPLEMENT FACTOR 3B
 CC TO GENERATE THE C3 OR C5 CONVERTASE.
 CC -1- CATALYTIC ACTIVITY: Cleaves C3 in the alpha-chain to yield C3a and
 CC C3b. Cleaves C5 in the alpha-chain to yield C5a and C5b. Both
 CC cleavages take place at the carboxinal of an arginine residue
 CC -1- MISCELLANEOUS: FACTOR B IS A MAJOR HISTOCOMPATIBILITY COMPLEX
 CC CLASS-II PROTEIN.
 CC -1- SIMILARITY: WITH COMPLEMENT C2.
 CC -1- SIMILARITY: CONTAINS 3 SUSHI (SC) DOMAINS
 CC -1- SIMILARITY: CONTAINS 1 WFPA DOMAIN.
 CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY SL; ALSO KNOWN AS THE
 CC TRYPSIN FAMILY.
 CC this SWISS PROT entry is copyright, it is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL collaboration
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb.scrib.com>)
 CC or send an email to license@isb.scrib.com.
 DR EMBL: M60646; AAA37379.1;
 DR EMBL: M60629; AAA37379.1; JOINED
 DR EMBL: M60640; AAA37379.1; JOINED
 DR EMBL: M60611; AAA37379.1; JOINED
 DR EMBL: M60612; AAA37379.1; JOINED
 DR EMBL: M60613; AAA37379.1; JOINED
 DR EMBL: M60614; AAA37379.1; JOINED
 DR EMBL: M60645; AAA37379.1; JOINED
 DR EMBL: M60646; AAA37379.1; JOINED
 DR EMBL: M60647; AAA37379.1; JOINED
 DR EMBL: M60648; AAA37379.1; JOINED
 DR EMBL: M60649; AAA37379.1; JOINED
 DR EMBL: M60640; AAA37379.1; JOINED
 DR EMBL: M60641; AAA37379.1; JOINED
 DR EMBL: M60642; AAA37379.1; JOINED
 DR EMBL: M60643; AAA37379.1; JOINED
 DR EMBL: M60644; AAA37379.1; JOINED
 DR EMBL: M60645; AAA37379.1; JOINED
 DR EMBL: M60646; AAA37379.1; JOINED
 DR EMBL: A1049850; AAC05283.1;

```

16 FMHL AF109906: AACGATGG 17
16 EMHL K01496: AAAAGGAG 12
16 EMHL K01497: AAAAGGAG 12
16 EMHL K01498: AAAAGGAG 12
16 EMHL M57690: AAAAGGAG 12
16 FMHL A00875: HMM
16 HSPD P00743: 2HNE
16 MEMPS S01196:
16 MDL MGI105979: H2 B1
16 InterPro: IPR001404: Chemotaxis
16 InterPro: IPR000406: Sushi SCR CCP
16 InterPro: IPR001254: LYPIN
16 InterPro: IPR002033: VWFA
16 Pfam: PF00084: sushi 1
16 Pfam: PF00089: LYPIN 2
16 Pfam: PF00092: vwa 1
16 PRINTS: PR00222: CHYMOTRYPSIN
16 PRINSE: PR00451: VWFA/MAIN
16 SMART: SM00042: CCP 3
16 SMART: SM00020: LYPIN 1
16 SMART: SM00427: VWA 1
16 PROSITE: PS00240: TRYPSIN DOM 1
16 PROSITE: PS00144: TRYPSIN HIS 1
16 PROSITE: PS00145: TRYPSIN SER 1
16 PROSITE: PS00244: VWFA 1
16 Complement alternative pathway: Plasma: Hydrolase: Serine protease
16 Glycoprotein Repeat / Sushi:
16 SIGNAL 1 22 BY SIMILARITY
16 CHAIN 24 761 COMPLEMENT FACTOR B
16 CHAIN 24 256 BA FRAGMENT
16 CHAIN 257 761 BA FRAGMENT
16 DOMAIN 34 96 SUSHI 1
16 DOMAIN 99 156 SUSHI 2
16 DOMAIN 161 216 SUSHI 3
16 DOMAIN 267 466 VWFA
16 DOMAIN 479 761 SERINE PROTEASE
16 AC1 SITE 523 523 CHARGE RELAY SYSTEM
16 AC1 SITE 574 574 CHARGE RELAY SYSTEM
16 AC1 SITE 696 696 CHARGE RELAY SYSTEM
16 DISULFID 34 74 BY SIMILARITY
16 DISULFID 59 56 BY SIMILARITY
16 DISULFID 100 142 BY SIMILARITY
16 DISULFID 128 155 BY SIMILARITY
16 DISULFID 162 202 BY SIMILARITY
16 DISULFID 188 216 BY SIMILARITY
16 CARBOHYD 119 119 N LINKED (GLNAC...) (POTENTIAL)
16 CARBOHYD 139 139 N LINKED (GLNAC...) (POTENTIAL)
16 CARBOHYD 282 282 N LINKED (GLNAC...) (POTENTIAL)
16 CARBOHYD 476 476 N LINKED (GLNAC...) (POTENTIAL)
16 CONFLICT 730 730 I - P (IN REF 4)
16 CONFLICT 745 745 V - E (IN REF 4)
16 SEQUENCE 761 AA: 80004 MW: 990154486.1020E CR'64

```


Copyright © 1997 by Greenwood Publishing Group, Inc.

IM protein protein search, using sw model

November 6, 2002, 16:05:00 ; Search time 21.4201 Seconds
(without alignments)
1082.224 Million cell updates/sec

US 09-834 309 A

total score: 756

Sequence: 1 G T C C G P P I L N G R I S Y S T A N M W C P T R L P T C V S V F P L F 134

Source input table: BL05SIM62

Gapex 10.0, Gapext 0.5

Searched: 562212 seqs, 17294929 residues

Total number of bits satisfying condition: 593677

Minimum IN seq length: 0

```
maximum length: 20000000
```

Postscript: Minimum (8)

Maximum Match 100%

Listing first 100 summaries

Database : SPIRIT:MH 19:*

- 1: sp_archaea: *
- 2: sp_bacteria: *
- 3: sp_fungi: *
- 4: sp_human: *
- 5: sp_invertebrate: *
- 6: sp_mammal: *
- 7: sp_monoc: *
- 8: sp_orchidlike: *
- 9: sp_plant: *
- 10: sp_protist: *
- 11: sp_rhodent: *
- 12: sp_virus: *
- 13: sp_vertebrate: *
- 14: sp_unclassified: *
- 15: sp_virus: *
- 16: sp_bacteriap: *
- 17: sp_archaeap: *

Prod. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

Result No.	Score	Query Match	Length	DB	ID	Description
1	751	100.0	1032	4	Q14866	Q14866 homo sapien
2	751	100.0	1087	4	Q14212	Q14212 homo sapien
3	466.5	62.1	679	11	Q98254	Q98254 mus musculus
4	466.5	62.1	1032	11	Q98884	Q98884 ovis aries
5	464.5	61.9	1045	6	Q66345	Q66345 papio hamad
6	235	41	1941	4	Q29528	Q29528 papio hamad
7	220	29.3	2039	4	Q16745	Q16745 homo sapien
8	220	29.3	2489	4	Q16744	Q16744 homo sapien
9	218	29.0	132	4	Q98899	Q98899 homo sapien
10	218	29.0	132	4	Q98898	Q98898 homo sapien
11	218	29.0	2034	6	Q29530	Q29530 pan troglod
12	215	28.6	132	4	Q98555	Q98555 homo sapien
13	215	28.6	259	12	H87616	H87616 cowpox viru
14	211	28.1	132	4	Q9854	Q9854 homo sapien
15	210	28.0	263	12	Q89859	Q89859 variola vir
16	210	28.0	263	12	Q07033	Q07033 variola vir


```

11 Q2962W PRELIMINARY: PRI: 1911 AA.
12 Q2962W
13 01 NOV 1996 (TREMREL: 01, Created)
14 01 NOV 1996 (TREMREL: 01, Last sequence update)
15 01 DEC 2001 (TREMREL: 19, Last annotation update)
16 COMPLEMENT COMPONENT RECEPTOR TYPE 1 (FRAGMENT).
17
18 05 Papio hamadryas (Hamadryas baboon).
19 06 Eukaryota: Metazoa: Chordata: Craniata: Vertebrata: Euteleostomi:
20 07 Mammalia: Eutheria: Primates: Catarrhini: Cercopithecoidea:
21 08 Cercopithecoidea: Papio.
22 09 NCBI TaxID: 9557.
23
24 RN
25
26 SEQUENCE FROM N.A.
27 RA Goldman L., Subramanian B.V., Nickells M.W., Bourcade D.E.,
28 Alkinson J.P.
29 *Primary sequence of the baboon 200 kDa C5b/C4b receptor (CRL).
30
31 01 Submitted (MAR 1995) to the EMBL/GenBank/TrEMBL databases.
32
33 DR EMBL: L39911 AAA62170.1;
34 DR EMBL: L39911 AAA62170.1;
35 DR BSSP: P08603; 1HET.
36 InterPro: IPR001424; S04C02 ZN.
37 InterPro: IPR000436; Sushi_SCR_CCP.
38 InterPro: IPR000844; Zn_carbopept.
39 Pfam: PF00084; Sushi: 29.
40 SMART: SM00042; CCP: 29.
41 PROSITE: PS00134; CARBOXYPEPT_2N_2; UNKNWN_4.
42 PROSITE: PS00087; S04C02 ZN_1; UNKNWN_1.
43 KW Receptor.
44
45 IT NCBI TER 1911 1911
46 SEQUENCE 1911 AA: 216174 MW: 555440DDAAFA52D CRC64;
47
48 Query Match 41.4%; Score 245; DR 6; Length 1911;
49 Best Local Similarity 48.4%; Pred. No. 5,26 17;
50 Matches 94; Conservative 16; Mismatches 54; Indels 16; Gaps 6;
51
52 QY 2 ISGSGSPPTNGRISYYS--TPVAVGVIRYSC SGTFRGLGKSLGDTIKKVV 52
53 111 111 111 111 111 111 111 111 111 111 111 111 111 111
54 1304 ISGKPPPTNSG DYSNNKISFSGGVVYQCHGCHKQKQFELVGRSVYVTSKDDQ 1461
55 111 111 111 111 111 111 111 111 111 111 111 111 111 111
56 53 DGTWVFATKGFVYSPYS*FEIVGSGYFEPSTF YPDQSVTATATNESMKNKSV 111
57 111 111 111 111 111 111 111 111 111 111 111 111 111 111
58 1362 VGWSSPPPRCTSNK--TAPEVKNATRVNKRFSFELIVGRQGVVWGSHIV 1518
59 111 111 111 111 111 111 111 111 111 111 111 111 111 111
60 112 WCUANMMWPIRLPCTVSV 140
61 111 111 111 111 111 111 111 111 111 111 111 111 111 111
62 1519 QUTNRWHP-KLPDPSRV 1649
63
64 RESULT 7
65 Q16745 PRELIMINARY: PRI: 2049 AA.
66 AC Q16745
67 01 NOV 1996 (TREMREL: 01, Created)
68 01 NOV 1996 (TREMREL: 01, Last sequence update)
69 01 DEC 2001 (TREMREL: 19, Last annotation update)
70 COMPLEMENT RECEPTOR 1.
71 Homo sapiens (Human).
72 Eukaryota: Metazoa: Chordata: Craniata: Vertebrata: Euteleostomi:
73 Mammalia: Eutheria: Primates: Catarrhini: Homiidae: Homo.
74 NCBI TaxID: 9606.
75
76 RN
77
78 SEQUENCE FROM N.A.
79 Vukobratovic W., Wink W.W.
80 Submitted (JUN 1994) to the EMBL/GenBank/TrEMBL databases.
81
82 DR EMBL: L17418; AAB60694.1;
83 DR EMBL: L17418; AAB60694.1; JOINED.

```

```

DR EMBL: L17499; AAB60694.1; JOINED.
DR EMBL: L17499; AAB60694.1; JOINED.
DR EMBL: L17419; AAB60694.1; JOINED.
DR EMBL: L17420; AAB60694.1; JOINED.
DR EMBL: L17421; AAB60694.1; JOINED.
DR EMBL: L17422; AAB60694.1; JOINED.
DR EMBL: L17423; AAB60694.1; JOINED.
DR EMBL: L17424; AAB60694.1; JOINED.
DR EMBL: L17425; AAB60694.1; JOINED.
DR EMBL: L17426; AAB60694.1; JOINED.
DR EMBL: L17427; AAB60694.1; JOINED.
DR EMBL: L17428; AAB60694.1; JOINED.
DR EMBL: L17429; AAB60694.1; JOINED.
DR EMBL: L17430; AAB60694.1; JOINED.
DR EMBL: L17431; AAB60694.1; JOINED.
DR EMBL: L17432; AAB60694.1; JOINED.
DR EMBL: L17433; AAB60694.1; JOINED.
DR EMBL: L17434; AAB60694.1; JOINED.
DR EMBL: L17435; AAB60694.1; JOINED.
DR EMBL: L17436; AAB60694.1; JOINED.
DR EMBL: L17437; AAB60694.1; JOINED.
DR BSSP: P08603; 1HET.
DR InterPro: IPR001424; S04C02 ZN.
DR InterPro: IPR000436; Sushi_SCR_CCP.
DR InterPro: IPR000844; Zn_carbopept.
DR Pfam: PF00084; Sushi: 30.
DR SMART: SM00042; CCP: 30.
DR PROSITE: PS00134; CARBOXYPEPT_2N_2; UNKNWN_2.
DR PROSITE: PS00087; S04C02 ZN_1; UNKNWN_1.
DR KW Receptor.
DR SEQUENCE 2049 AA: 224617 MW: 108267116H16635 CRC64;
49
50 Query Match 29.4%; Score 229; DR 4; Length 2049;
51 Best Local Similarity 47.4%; Pred. No. 2,64 15;
52 Matches 52; Conservative 16; Mismatches 55; Indels 16; Gaps 6;
53
54 QY 2 ISGSGSPPTNGRISYYS--TPVAVGVIRYSC SGTFRGLGKSLGDTIKKVV 52
55 111 111 111 111 111 111 111 111 111 111 111 111 111 111
56 1517 ISGKPPPTNSG DYSNNKISFSGGVVYQCHGCHKQKQFELVGRSVYVTSKDDQ 1574
57 111 111 111 111 111 111 111 111 111 111 111 111 111 111
58 53 DGTWVFATKGFVYSPYS*FEIVGSGYFEPSTF YPDQSVTATATNESMKNKSV 111
59 111 111 111 111 111 111 111 111 111 111 111 111 111 111
60 1575 VGWSSPPPRCTSNK--TAPEVKNATRVNKRFSFELIVGRQGVVWGSHIV 1641
61 111 111 111 111 111 111 111 111 111 111 111 111 111 111
62 112 WCUANMMWPIRLPCTVSV 140
63 111 111 111 111 111 111 111 111 111 111 111 111 111 111
64 1632 QUTNRWHP-KLPDPSRV 1649
65
66 RESULT 8
67 Q16744 PRELIMINARY: PRI: 2489 AA.
68 AC Q16744
69 01 NOV 1996 (TREMREL: 01, Created)
70 01 NOV 1996 (TREMREL: 01, Last sequence update)
71 01 DEC 2001 (TREMREL: 19, Last annotation update)
72 COMPLEMENT RECEPTOR 1.
73 Homo sapiens (Human).
74 Eukaryota: Metazoa: Chordata: Craniata: Vertebrata: Euteleostomi:
75 Mammalia: Eutheria: Primates: Catarrhini: Homiidae: Homo.
76 NCBI TaxID: 9606.
77
78 RN
79
80 SEQUENCE FROM N.A.
81 Vukobratovic W., Wink W.W.
82 Submitted (JUN 1994) to the EMBL/GenBank/TrEMBL databases.
83
84 DR EMBL: L17418; AAB60694.1;
85 DR EMBL: L17418; AAB60694.1; JOINED.

```

SEQUENCE FROM N.A.
MEDLINE:94065175; PubMed:8245464;

Vik D.P., Wong W.W.

Structure of the gene for the E allele of complemented receptor type 1 and sequence of the coding region unique to the S allele.

J. Immunol. 151:6214-6224(1993).

[2]

SEQUENCE FROM N.A.

Vik D.P., Wong W.W.

Submitted (JUN 1994) to the EMBL/GenBank/DBJ databases

EMBL: L17418; AAB60695.1; JOINED.

EMBL: L17400; AAB60695.1; JOINED.

EMBL: L17409; AAB60695.1; JOINED.

EMBL: L17419; AAB60695.1; JOINED.

EMBL: L17420; AAB60695.1; JOINED.

EMBL: L17421; AAB60695.1; JOINED.

EMBL: L17422; AAB60695.1; JOINED.

EMBL: L17423; AAB60695.1; JOINED.

EMBL: L17424; AAB60695.1; JOINED.

EMBL: L17425; AAB60695.1; JOINED.

EMBL: L17426; AAB60695.1; JOINED.

EMBL: L17427; AAB60695.1; JOINED.

EMBL: L17428; AAB60695.1; JOINED.

EMBL: L17429; AAB60695.1; JOINED.

EMBL: L17430; AAB60695.1; JOINED.

EMBL: L17431; AAB60695.1; JOINED.

EMBL: L17432; AAB60695.1; JOINED.

EMBL: L17433; AAB60695.1; JOINED.

EMBL: L17434; AAB60695.1; JOINED.

EMBL: L17435; AAB60695.1; JOINED.

EMBL: L17436; AAB60695.1; JOINED.

EMBL: L17437; AAB60695.1; JOINED.

EMBL: L17438; AAB60695.1; JOINED.

EMBL: L17439; AAB60695.1; JOINED.

EMBL: L17440; AAB60695.1; JOINED.

EMBL: L17441; AAB60695.1; JOINED.

EMBL: L17442; AAB60695.1; JOINED.

EMBL: L17443; AAB60695.1; JOINED.

EMBL: L17444; AAB60695.1; JOINED.

EMBL: L17445; AAB60695.1; JOINED.

EMBL: L17446; AAB60695.1; JOINED.

EMBL: L17447; AAB60695.1; JOINED.

EMBL: L17448; AAB60695.1; JOINED.

EMBL: L17449; AAB60695.1; JOINED.

EMBL: L17450; AAB60695.1; JOINED.

EMBL: L17451; AAB60695.1; JOINED.

EMBL: L17452; AAB60695.1; JOINED.

EMBL: L17453; AAB60695.1; JOINED.

EMBL: L17454; AAB60695.1; JOINED.

EMBL: L17455; AAB60695.1; JOINED.

EMBL: L17456; AAB60695.1; JOINED.

EMBL: L17457; AAB60695.1; JOINED.

EMBL: L17458; AAB60695.1; JOINED.

EMBL: L17459; AAB60695.1; JOINED.

EMBL: L17460; AAB60695.1; JOINED.

EMBL: L17461; AAB60695.1; JOINED.

EMBL: L17462; AAB60695.1; JOINED.

EMBL: L17463; AAB60695.1; JOINED.

EMBL: L17464; AAB60695.1; JOINED.

EMBL: L17465; AAB60695.1; JOINED.

EMBL: L17466; AAB60695.1; JOINED.

EMBL: L17467; AAB60695.1; JOINED.

EMBL: L17468; AAB60695.1; JOINED.

EMBL: L17469; AAB60695.1; JOINED.

EMBL: L17470; AAB60695.1; JOINED.

EMBL: L17471; AAB60695.1; JOINED.

EMBL: L17472; AAB60695.1; JOINED.

EMBL: L17473; AAB60695.1; JOINED.

EMBL: L17474; AAB60695.1; JOINED.

EMBL: L17475; AAB60695.1; JOINED.

EMBL: L17476; AAB60695.1; JOINED.

EMBL: L17477; AAB60695.1; JOINED.

EMBL: L17478; AAB60695.1; JOINED.

EMBL: L17479; AAB60695.1; JOINED.

EMBL: L17480; AAB60695.1; JOINED.

EMBL: L17481; AAB60695.1; JOINED.

EMBL: L17482; AAB60695.1; JOINED.

EMBL: L17483; AAB60695.1; JOINED.

EMBL: L17484; AAB60695.1; JOINED.

EMBL: L17485; AAB60695.1; JOINED.

EMBL: L17486; AAB60695.1; JOINED.

QY 112 WCQANNMGGTTRLPQCVS 130

DB 2082 QYQIN:pwat-KLPQSRV 2099

RESULT 9

Q9HB98

ID Q9HB98 PRELIMINARY; PRT: 132 AA.

AC Q9HB98

DT 01-MAR-2001 (Tremblrel. 16, Created)

DT 01-MAR-2001 (Tremblrel. 16, Last sequence update)

DE 01-DEC-2001 (Tremblrel. 19, Last annotation update)

GN CRI

OS Homo sapiens (human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi.

OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

OX NCB:TaxID:9606;

RN 111

RP Zimmerman P.A., McNamara D.T., Miller L.L., Doubo O., Moulds J.

RL Submitted (MAY 2000) to the EMBL/GenBank/DBJ databases.

DR EMBL: AF264715; AAG14442.1; -

DR HSP: P02749; I00H.

DR InterPro: IPR000436; Sushi_SCR_CCP.

DR Pfam: PF00084; sushi; 2.

DR SMART: SM00032; CCP; 2.

KW Receptor.

FT NON_TER 1 1

FT VARIANT 94 94 1 -> S.

FT NON_TER 132 132

SQ SEQUENCE 132 AA; 14680 MW; 71A9B91B5596958B CRC64;

Query Match 29.0%; Score 218; DB 4; Length 132;

Best local Similarity 36.8%; Pred. No. 156-15;

Matches 50; Conservative 17; Mismatches 54; Indels 16; Gaps 6;

QY 2 ISGSPPTLNGRISYYS---TPAVGVIRASQ-----SCTFELIGKSLCTTKKV 52

DB 1 ISCEPPPTISNG-DFYNNRTSFUNGTVVTVYQHIGPDCGQLFELVGEISYCTSKDDQ 58

QY 53 DCTWDKPAKCEYFNKYSSCPPEPVPQYKIPGSP--YHHSVTFACKTNFSMNKNSV 111

DB 59 VGVWSSPPPCISINK--CTAPEVENA:RVFGNKFETLILIRKQGVVMVSHIV 119

QY 112 WCQANNMGGTTRLPQCVS 130

DB 116 QCYTNGRQI-KLPQC 130

RESULT 10

Q9HB98

ID Q9HB98 PRELIMINARY; PRT: 132 AA.

AC Q9HB98

DT 01-MAR-2001 (Tremblrel. 16, Created)

DT 01-MAR-2001 (Tremblrel. 16, Last sequence update)

DE 01-DEC-2001 (Tremblrel. 19, Last annotation update)

GN CRI

OS Homo sapiens (human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi.

OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

OX NCB:TaxID:9606;

RN 111

RP Zimmerman P.A., McNamara D.T., Miller L.L., Doubo O., Moulds J.

RL Submitted (MAY 2000) to the EMBL/GenBank/DBJ databases.

DR EMBL: AF264715; AAG14443.1; -

DR HSP: P10998; IVD.

DR InterPro: IPR000436; Sushi_SCR_CCP.

DR Pfam: PF00084; sushi; 2.

DR SMART: SM00032; CCP; 2.

QY 2 ISGSPPTLNGRISYYS---TPAVGVIRASQ-----SCTFELIGKSLCTTKKV 52

DB 1967 ISCEPPPTISNG-DFYNNRTSFUNGTVVTVYQHIGPDCGQLFELVGEISYCTSKDDQ 2024

QY 53 DCTWDKPAKCEYFNKYSSCPPEPVPQYKIPGSP--YHHSVTFACKTNFSMNKNSV 111

DB 2025 VGVWSSPPPCISINK--CTAPEVENA:RVFGNKFETLILIRKQGVVMVSHIV 2081

Query Match 29.0%; Score 220; DB 4; Length 2489;

Best local Similarity 37.4%; Pred. No. 33e-15;

Matches 52; Conservative 16; Mismatches 55; Indels 16; Gaps 6;

QY 2 ISGSPPTLNGRISYYS---TPAVGVIRASQ-----SCTFELIGKSLCTTKKV 52

DB 1967 ISCEPPPTISNG-DFYNNRTSFUNGTVVTVYQHIGPDCGQLFELVGEISYCTSKDDQ 2024

QY 53 DCTWDKPAKCEYFNKYSSCPPEPVPQYKIPGSP--YHHSVTFACKTNFSMNKNSV 111

DB 2025 VGVWSSPPPCISINK--CTAPEVENA:RVFGNKFETLILIRKQGVVMVSHIV 2081

RA Ryazankina O.I., Gutorov V.V., Kotwal G.J.:
 RT *The genomic sequence analysis of the left and right species-specific
 RC terminal region of a cowpox virus strain reveals unique sequences and
 RD a cluster of intact ORFs for immunomodulatory and host range
 RE proteins.*

RL Virology 24:432-460(1998).
 DR EMBL: X94355; CAA64102.1;
 DR EMBL: Y11842; CAA72567.1;
 DR HSSP: P10998; LVVD.
 DR InterPro: IPR000436; Sushi_SRF_CCP.
 DR Pfam: PF00084; sushi; 4.
 DR SMART: SM00042; CCP; 4.
 SQ SEQUENCE 259 AA: 28193 MW: 901AAEF6894BR59A CRC64;

Query Match 28.6%; Score 215; DR 12; Length 259;
 Best Local Similarity 35.9%; Pred. No. 7.7e-16;
 Matches 46; Conservative 15; Mismatches 55; Indels 12; Gaps 5;

QY 2 ISGSPPTILNGKISVYSTPIAVGVIVYSGTFLIGKSLDITKDKVDFWKPAP 61
 DR 142 VKQSPHSISNGKINGCYNYDFYDGSVVYVYSGSYSLIGNSCVLC ----SGGWSDP-P 195
 QY 62 KTEYFNKYSSCPPEIVPGYKIRG SDYPRGDSVTFACKTNFMSNKNKSVWQANMWG 120
 DR 196 TCGIV ---KCPPTILNGYSSGPKRVSYNVDKCKYKISNSSSTSPSPNTWQ 251
 QY 121 PTRLPTCV 128
 DR 252 P-ELPCV 258

RESULT 14
 QYHC54
 ID QYHC54 PRELIMINARY; PRI; 132 AA.
 AC QYHC54;
 DT 01-MAR-2001 (TREMblrel, 16, Created)
 DT 01-MAR-2001 (TREMblrel, 16, Last sequence update)
 DT 01-DEC-2001 (TREMblrel, 19, Last annotation update)
 DE COMPLEMENT RECEPTOR 1 (FRAGMENT).
 GN CR1.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
 CX NCBI TaxID: 9606;
 FN 111

PP SEQUENCE FROM N.A.
 RA Zimmerman P.A., McNamara D.J., Birmingham D., Atkinson J.P.,
 FA Miller L., Doubo O., Moulds J.,
 RL Submitted (JUL-1999) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AF169970; AAC09140.1;
 DR HSSP: P10998; LVVD.
 DR InterPro: IPR000436; Sushi_SRF_CCP.
 DR Pfam: PF00084; sushi; 2.
 DR SMART: SM00042; CCP; 2.
 KW Receptor.
 FT NON_TPR 1 1
 FT VARIANT 74 74 E -> K.
 FT VARIANT 85 85 G -> R.
 FT NON_TPR 132 132
 SQ SEQUENCE 132 AA: 14554 MW: 1AD1159EP5989FCB CRC64;

Query Match 28.1%; Score 211; DR 4; Length 132;
 Best Local Similarity 46.8%; Pred. No. 9.5e-16;
 Matches 50; Conservative 16; Mismatches 54; Indels 16; Gaps 6;

QY 2 ISGSPPTILNGKISVYSGTPIAVGVIVYSGTFLIGKSLDITKDKV 52
 DR 142 VKQSPHSISNGKINGCYNYDFYDGSVVYVYSGSYSLIGNSCVLC ----SGGWSDP-P 195
 QY 62 KTEYFNKYSSCPPEIVPGYKIRG SDYPRGDSVTFACKTNFMSNKNKSVWQANMWG 120
 DR 196 TCGIV ---KCPPTILNGYSSGPKRVSYNVDKCKYKISNSSSTSPSPNTWQ 251
 QY 121 PTRLPTCV 128
 DR 252 P-ELPCV 258

QY 112 WCQANMMWGPTRLPTC 127
 DR 116 QCQTNGWGP-KLPWC 130
 RESULT 15
 QYHC54
 ID QYHC54 PRELIMINARY; PRI; 263 AA.
 AC QYHC54;
 DT 01-NOV-1996 (TREMblrel, 01, Created)
 DT 01-NOV-1996 (TREMblrel, 01, Last sequence update)
 DT 01-JUN-2001 (TREMblrel, 17, Last annotation update)
 DE BOWLING OF VACCINIA VIRUS CDS C31
 GN D15L.
 OS Variola virus.
 OC Viruses; dsDNA viruses, no RNA stage; Poxviridae; Chordopoxvirinae;
 CC Orthopoxvirus.
 CX NCBI TaxID: 10255;
 FN 111

PP SEQUENCE FROM N.A.
 RA Massung F.F., Esposito J.J., Liu S.L., Qi L., Griebel L.K.,
 RA Knight J.C., Aubin L., Yuran T.E., Parsons T.M., Loparev V.N.,
 RT *potential virulence determinants in terminal regions of variola
 RT smallpox virus genome.*
 RL Nature 366:748-751(1993).
 RN 121
 RP SEQUENCE FROM N.A.
 RC STRAIN-SMALLA-1977;
 RA Massung F.F., Loparev V.N., Knight J.C., Chizhikov V.E., Parsons T.M.,
 RA Totmenko A.V., Shchelkunov S.N., Esposito J.J.;
 RL Submitted (JUL-1995) to the EMBL/GenBank/DBJ databases.
 DR EMBL: 122573; AAA60760.1;
 DR EMBL: U1834; AAA69423.1;
 DR HSSP: P10998; LVVD.
 DR InterPro: IPR000436; Sushi_SRF_CCP.
 DR Pfam: PF00084; sushi; 4.
 DR SMART: SM00042; CCP; 4.
 SQ SEQUENCE 263 AA: 28816 MW: 046597800294220 CRC64;

Query Match 28.0%; Score 20; DR 12; Length 264;
 Best Local Similarity 35.2%; Pred. No. 2.8e-15;
 Matches 41; Conservative 16; Mismatches 55; Indels 12; Gaps 5;

QY 2 ISGSPPTILNGKISVYSTPIAVGVIVYSGTFLIGKSLDITKDKVDFWKPAP 61
 DR 146 VKQSPHSISNGKINGCYNYDFYDGSVVYVYSGSYSLIGNSCVLC ----SGGWSDP-P 199
 QY 62 KTEYFNKYSSCPPEIVPGYKIRG SDYPRGDSVTFACKTNFMSNKNKSVWQANMWG 120
 DR 200 TCGIV ---KCPPTILNGYSSGPKRVSYNVDKCKYKISNSSSTSPSPNTWQ 255
 QY 121 PTRLPTCV 128
 DR 256 P-ELPCV 262

RESULT 16
 QYHC54
 ID QYHC54 PRELIMINARY; PRI; 263 AA.
 AC QYHC54;
 DT 01-NOV-1996 (TREMblrel, 01, Created)
 DT 01-NOV-1996 (TREMblrel, 01, Last sequence update)
 DT 01-JUN-2001 (TREMblrel, 17, Last annotation update)
 DE D12L PROTEIN.
 GN D12L.
 OS Variola virus.
 OC Viruses; dsDNA viruses, no RNA stage; Poxviridae; Chordopoxvirinae;
 CC Orthopoxvirus.
 CX NCBI TaxID: 10255;
 FN 111

PP SEQUENCE FROM N.A.

```

60 STRAIN INDIA 1967, R.F. MAJORS.
61 MEDLINE 94302281; PubMed 846129;
62 Chizhikov V.N., Loparev V.V., Kulzhik V.V.,
63 "Genes of variola and vaccinia viruses necessary to overcome the host
64 protective mechanisms."
65 PNAS 94, 419-420 (1994).
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
974
975
976
977
978
979
980
981
982
983
984
985
986
987
988
989
990
991
992
993
994
995
996
997
998
999
1000

```

```

60 STRAIN INDIA 1967, R.F. MAJORS.
61 MEDLINE 94302281; PubMed 846129;
62 Chizhikov V.N., Loparev V.V., Kulzhik V.V.,
63 "Genes of variola and vaccinia viruses necessary to overcome the host
64 protective mechanisms."
65 PNAS 94, 419-420 (1994).
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
974
975
976
977
978
979
980
981
982
983
984
985
986
987
988
989
990
991
992
993
994
995
996
997
998
999
1000

```

```

Query Match 27.78; Score 208; DB 12; Length 264;
Best Local Similarity 35.28; Pred. No. 4.7e 15;
Matches 45; Conservative 16; Mismatches 55; Indels 12; Gaps 5;

CY 2 TSGSSPPFLNGRISYVSIPIAWGVIRKVSSTETPLGFKSLGTFKRWKWTWKAP 61
DB 146 VKQLPISINGRHNGYNDYTRISVVTYSNCSYSLICNSGVLC -- SGTWNSNP P 199
CY 52 KCVIRKVSSTETPLGFKSLGTFKRWKWTWKAP 120
DB 52 KCVIRKVSSTETPLGFKSLGTFKRWKWTWKAP 120
CY 200 PQGIV- KCPYPIINGVLSGSKRSYSYNQVIEPDKYVKLSGSSSTSGPNTWQ 255
DB 200 PQGIV- KCPYPIINGVLSGSKRSYSYNQVIEPDKYVKLSGSSSTSGPNTWQ 255
CY 121 PTKLPTGV 128
DB 121 PTKLPTGV 128
CY 256 P-ELPKIV 262
DB 256 P-ELPKIV 262

RESULT 18
Q89061 PRELIMINARY; PRI: 263 AA.
AC Q89061;
DT 01 NOV 1996 (TEMBREL 01, Created)
ID 01 NOV 1996 (TEMBREL 01, Last sequence update)
DE 01 DEC 2001 (TEMBREL 19, Last annotation update)
DE 0151
OS Variola virus.
OC Viruses; dsDNA viruses, no RNA stage; Poxviridae; Chordopoxvirinae;
OC Orthopoxvirus.
OX NCBI_TaxID=10255;
RN [1]
RC SEQUENCE FROM N.A.
RX MEDLINE=96295428; PubMed=8661449;
RA Massum R.F., Loparev V.N., Kulzhik V.V., Tolmenin A.V.,
RA Chizhikov V.E., Parsons J.M., Safonov P.F., Gurorov V.V.,
RA Shchelkunov S.N., Esposito J.J.,
RT "Terminal region sequence variations in variola virus DNA."
PL Virology 221:291-300(1996);
DR EMOL: 018347; AAA69317.1; -.
DR BSSP: P10998; LVVD.
DR InterPro: IPR000436; Sushi_SCR_CCP.
DR Pfam: PF00084; Sushi; 4.
DR SMART: SM00042; CCP; 4.
SQ SEQUENCE 263 AA; 28843 MW; 04509964FAV4220 CR664;

Query Match 27.78; Score 208; DB 12; Length 264;
Best Local Similarity 35.28; Pred. No. 4.7e 15;
Matches 45; Conservative 16; Mismatches 55; Indels 12; Gaps 5;

CY 2 TSGSSPPFLNGRISYVSIPIAWGVIRKVSSTETPLGFKSLGTFKRWKWTWKAP 61
DB 146 VKQLPISINGRHNGYNDYTRISVVTYSNCSYSLICNSGVLC -- SGTWNSNP P 199
CY 52 KCVIRKVSSTETPLGFKSLGTFKRWKWTWKAP 120
DB 52 KCVIRKVSSTETPLGFKSLGTFKRWKWTWKAP 120
CY 200 PQGIV- KCPYPIINGVLSGSKRSYSYNQVIEPDKYVKLSGSSSTSGPNTWQ 255
DB 200 PQGIV- KCPYPIINGVLSGSKRSYSYNQVIEPDKYVKLSGSSSTSGPNTWQ 255
CY 121 PTKLPTGV 128
DB 121 PTKLPTGV 128
CY 256 P-ELPKIV 262
DB 256 P-ELPKIV 262

```



```

SEQUENCE FROM N.A.
RC STRAIN-HARTLEY, TISSUE-TESTIS;
RA Hosokawa M., Nonaka M., Okada N., Nonaka M., Okada H.
R1 "Molecular cloning of guinea pig membrane cofactor
RT protein-preferential expression in testis."
RL J. Immunol. 0:0-0(1996).
RN 12
SEQUENCE FROM N.A.
RC STRAIN-HARTLEY, TISSUE-TESTIS;
RX MEDLINE-97098723; PubMed-8943400;
RA Hosokawa M., Nonaka M., Okada N., Nonaka M., Okada H.
R1 "Molecular cloning of guinea pig Membrane Cofactor
RT Protein-preferential expression in testis."
RL J. Immunol. 157:4946-4952(1996).
DR EMBL: D84140; BAA12242.1;
DR HSSP: P10998; LVVD;
DR InterPro: IPR000436; SUSHI_SCR_COP;
DR EMBL: P10984; SUSHI; 4;
DR SMART: SM00342; COP; 4;
KW Signal.
FI NON_TER 1 19
FI SIGNAL 1 1
FI CHAIN 20 339
FI CHAIN MEMBRANE COFACTOR PROTEIN (CMI).
SQ SEQUENCE 339 AA: 37870 MW: 14610076 Da; CR64;

Query Match 22.8% Score 171; DB 11; Length 492;
Best local similarity 29.9% Pred. No. 66;
Matches 40; Conservative 21; Mismatches 57; Indels 16; Gaps 6;

QY 2 ISGSEPPPTINGKISYSTPT AVGVIVKYSNGT- ERLDERSLTCTTKRWVACT 55
Db 145 VLSPPKTLQNGKYFSSDVLVFEYFEAVTYS DAVGSHTRSLVGNFVYAGAGHGR 101
QY 56 WDKFAIKCEYFNKYSCPTVAGGVKRG STPYRHGDSVFAKNSMNKNSVWQ 114
Db 101 FEB-1997 (EMBLrel. 02, Last sequence update)
DT 01-FEB-1997 (EMBLrel. 02, Last annotation update)
DE MEMBRANE COFACTOR PROTEIN (CMT-FULL) PRECURSOR.
OS Cavia porcellus (Guinea pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Hystriocomphili; Caviidae; Cavia.
OX NCBI_TaxID:10141;
RN 11
SEQUENCE FROM N.A.
RC STRAIN-HARTLEY, TISSUE-TESTIS;
RA Hosokawa M., Nonaka M., Okada N., Nonaka M., Okada H.
R1 "Molecular cloning of guinea pig membrane cofactor protein
RT preferential expression in testis."
RL J. Immunol. 0:0-0(1996).
RN 12
SEQUENCE FROM N.A.
RC STRAIN-HARTLEY, TISSUE-TESTIS;
RX MEDLINE-97098723; PubMed-8943400;
RA Hosokawa M., Nonaka M., Okada N., Nonaka M., Okada H.
R1 "Molecular cloning of guinea pig Membrane Cofactor
RT Protein-preferential expression in testis."
RL J. Immunol. 157:4946-4952(1996).
DR EMBL: D84140; BAA12242.1;
DR HSSP: P10998; LVVD;
DR InterPro: IPR000436; SUSHI_SCR_COP;
DR EMBL: P10984; SUSHI; 4;
DR SMART: SM00342; COP; 4;
KW Signal.
FI NON_TER 1 19
FI SIGNAL 1 1
FI CHAIN 20 339
FI CHAIN MEMBRANE COFACTOR PROTEIN (CMI).
SQ SEQUENCE 339 AA: 37870 MW: 14610076 Da; CR64;

```

```

RESULT 42
P70105
ID P70105 PRELIMINARY; PRI: 4/9 AA.
AC P70105;
DT 01-FEB-1997 (EMBLrel. 02, Created)
DT 01-FEB-1997 (EMBLrel. 02, Last sequence update)
DT 01-DEC-2001 (EMBLrel. 19, Last annotation update)
DE MEMBRANE COFACTOR PROTEIN (CMT-FULL) PRECURSOR.
OS Cavia porcellus (Guinea pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Hystriocomphili; Caviidae; Cavia.
OX NCBI_TaxID:10141;
RN 11
SEQUENCE FROM N.A.
RC STRAIN-HARTLEY, TISSUE-TESTIS;
RA Hosokawa M., Nonaka M., Okada N., Nonaka M., Okada H.
R1 "Molecular cloning of guinea pig membrane cofactor protein
RT preferential expression in testis."
RL J. Immunol. 0:0-0(1996).
RN 12
SEQUENCE FROM N.A.
RC STRAIN-HARTLEY, TISSUE-TESTIS;
RX MEDLINE-97098723; PubMed-8943400;
RA Hosokawa M., Nonaka M., Okada N., Nonaka M., Okada H.
R1 "Molecular cloning of guinea pig Membrane Cofactor
RT Protein-preferential expression in testis."
RL J. Immunol. 157:4946-4952(1996).
DR EMBL: D84140; BAA12242.1;
DR HSSP: P10998; LVVD;
DR InterPro: IPR000436; SUSHI_SCR_COP;
DR EMBL: P10984; SUSHI; 4;

```



```

DR SMART: SM00042; CCP; 4.
KW SIGNAL. 1 34 POTENTIAL.
FT CHAIN 45 479 POTENTIAL.
SQ SEQUENCE 479 AA; 42159 MW; 50A56B058C8A1573 CRC64;

Query Match
Best Local Similarity 22.8%; Score 171; DB 11; Length 379;
Matches 40; Conservative 21; Mismatches 57; Indels 16; Gaps 6;

QY 2 ISGSPPEILNGKISYYSTPI AVGVIVIRYSSGT-----EFLIGERSLLCITKUKVDGT 55
DB 160 VLSSTPIKQNGKQIFSVQVFEFVAVSCDAVQCDKIKLSLGNFVLYCAGHOK----- 215
QY 56 WKPAPKIEYENKYSKSCPEPIVGGYKIRG-STPYRHHSVTFACIKTFNSMGNKNSVMQC 114
DB 216 WSSAAPE-----KVKKCP PVKNKQSLISGLGCTFFVATVTFQCLPGFYFNSSTVCC 271
QY 115 ANMMWPIRLPCV 128
DB 272 SNTWKPS-IPETL 284

RESULT 43
QYQAX6 PRELIMINARY; PRI: 268 AA.
AC QYQAX6;
DT 01 MAY 2000 (Tremblrel, 13, Created)
DT 01 MAY 2000 (Tremblrel, 13, Last sequence update)
DT 01 JUN 2001 (Tremblrel, 17, Last annotation update)
DB YH 05R.
DB YH 05R.
OS Yaba monkey tumor virus
OC Viruses; dsDNA viruses, no RNA stage; Poxviridae; Chordopoxvirinae;
OC Yatapoxvirus.
OX NCBI_TaxID=38804;
RN 111
RP SEQUENCE FROM N.A.
RA Anano H., Kato K., Miyamura T.;
RT *Sequence analysis of Yaba virus DNA.*;
RL Submitted (MAR-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL: AH025319; HA088977.1;
DR HSPD: P10908; LVVD.
DR InterPro: IPR000436; Sushi_SCR_CCP.
DR Pfam: PF00384; sushi; 3.
DR SMART: SM00042; CCP; 4.
SQ SEQUENCE 268 AA; 30606 MW; 4EF681CE5A94D6EC CRC64;

Query Match
Best Local Similarity 22.6%; Score 169.5; DB 12; Length 268;
Matches 41; Conservative 21; Mismatches 54; Indels 15; Gaps 5;

QY 3 SSGSPPEILNGKISYYSTPIAVGVIRYSSG-----GFLIGERSLLCITKDKVDGTWK 58
DB 79 NCKDIPVILNGNKNKQKYSLSGSLVYMKVKNKILERSLVAGNFTVKL-----DGKWP 134
QY 59 PAPKIEYEN-KYSSCPPEPIVGGYKIRG-STPYRHHSVTFACIKTFNSMGNKNSVMQANN 117
DB 134 DNEPCKLIRCKYFALQNHLEISPKKK-----FYGDVVIFKRTGFLSGSLTSGGINF 189
QY 118 MWCPRLPCV 128
DB 190 VWVP-NLPKV 199

RESULT 44
P88904 PRELIMINARY; PRI: 550 AA.
AC P88904;
DT 01 MAY 1997 (Tremblrel, 03, Created)
DT 01 MAY 1997 (Tremblrel, 03, Last sequence update)
DT 01 DEC 2001 (Tremblrel, 19, Last annotation update)
DE ORF 4.

OS Kaposi's sarcoma-associated herpesvirus (KSHV) (human herpesvirus 8)
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Gammaparvovirinae; Rhadinovirus.
OX NCBI_TaxID=37296;
RN 111
RP SEQUENCE FROM N.A.
RA Medline-97094384; PubMed-8939871;
RA Moore P.S., Hoshoff C., Weiss K.A., Chaud Y.;
RT *Molecular mimicry of human cytokine and cytokine response pathway genes by KSHV.*;
RL Science 274:1739-1744 (1996)
RN 121
RP SEQUENCE FROM N.A.
RA Russo J.J., Behenzky R.A., Chien M.-T., Chen J., Yan M., Maddalena D., Parry J.P., Peruzzi D., Edelman I.S., Chaud Y., Moore P.S.;
RA Parry J.P., Peruzzi D., Edelman I.S., Chaud Y., Moore P.S.;
RT *Nucleotide sequence of the Kaposi's Sarcoma associated herpesvirus (KSHV).*;
RL Proc. Natl. Acad. Sci. U.S.A. 93:14852-14867 (1996).
RN 131
RP SEQUENCE FROM N.A.
RA Russo J.J., Behenzky R.A., Chien M.-T., Chen J., Yan M., Maddalena D., Parry J.P., Peruzzi D., Edelman I.S., Chaud Y., Moore P.S.;
RL Submitted (MAY-1997) to the EMBL/GenBank/DBJ databases.
DR EMBL: U75996; AAC57082.1;
DR HSPD: P10997; LVVD.
DR InterPro: IPR000436; Sushi_SCR_CCP.
DR Pfam: PF03064; sushi; 4.
DR SMART: SM00042; CCP; 4.
SQ SEQUENCE 550 AA; 60688 MW; D488424A4AC1C85 CRC64;

Query Match
Best Local Similarity 22.6%; Score 169.5; DB 12; Length 550;
Matches 42; Conservative 14; Mismatches 51; Indels 19; Gaps 7;

QY 8 PPIINCR-----SYSTPIAVGVIRYSSGFLIGERSVTFACIKTFNSMGNKNSVMQANNWPT 63
DE 156 PKIKNGDKPKDKDYK-----YNDVAVHFFCNKYTLVGHPSIAC-----AVNNTWISNMP 207
QY 64 EYENKYSKSCPEPIVGGYKIRG-STPYRHHSVTFACIKTFNSMGNKNSVMQANNWPT 122
DB 208 E-----LACCKFPVSIVHGYPVQFSLIKYKIKSVTFACIKTFNSMGNKNSVMQANNWPT 262
QY 123 RLPTCV 128
DB 263 -LPKCV 267

RESULT 45
O40912 PRELIMINARY; PRI: 550 AA.
AC O40912;
DT 21-JAN-1998 (Tremblrel, 05, Created)
DT 21-JAN-1998 (Tremblrel, 05, Last sequence update)
DT 21-DEC-2001 (Tremblrel, 19, Last annotation update)
DE ORF 04.
OS Kaposi's sarcoma associated herpesvirus (KSHV) (human herpesvirus 8)
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Gammaparvovirinae; Rhadinovirus.
OX NCBI_TaxID=37296;
RN 111
RP SEQUENCE FROM N.A.
RA Medline-97296220; PubMed-9151804;
RA Neipel F., Albrecht J.C., Fleckenstein B.;
RT *Cell-homologous genes in the Kaposi's Sarcoma-associated rhadinovirus human herpesvirus 8: determinants of its pathogenicity?;
RL Virol. 71:4187-4192 (1997).
RN 121
RP SEQUENCE FROM N.A.
RA Neipel F., Albrecht J.C., Essner A., Huate Y.-Q., Li J.-J.;
RA Friedman-Kien A.E., Fleckenstein B.;
RT *The genome of human herpesvirus 8 cloned from Kaposi's Sarcoma.*;
RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.

```

```

106 1MB1: 090802: AAR02602.1:
106 R580: P10994: LVVH
106 InterPro: IPR00436: Sushi SCR_CYP
106 Pfam: PF00084: sushi; 4
106 SMART: SM0042: SCF; 4
106 SEQUENCE: 550 AA: 60648 MW: 347610 GF796A6 CRC64
Query Match: 2248; Score 108.5; DB 12; Length 550;
Best local similarity: 33.4%; Pred. No. 2.7e 10;
Matches 4; Conservative 14; Mismatches 51; Indels 19; Gaps 7;
CY 6 PFINDRI SYSLPTAVGLIVIKYSNGELRIGESLLGTLKDKVGLWELAPK 63
11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11
106 156 PFINDRIKDYEE YNDVHFENEGYTLVGPSTLAV--AVNNIWLNNMPT 207
CY 64 EYENKYSCTEPTAVGCKLRG STYRIGDSVTACKTFNSMGNKSNWQANNMNGPT 122
11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11
106 208 F LAGKFFSVLRKSYTQGLSLYKIKQSVTFACNGLFVLRGSPITLVNVEWDD 262
CY 133 RLPDY 128
11 11
106 263 LKDY 267

```

Search completed: November 6, 2002, 16:08:22
 Database: 27 4201 seqs

GenCore version 5.1.3
Copyright (c) 1994 - 2002 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 6, 2002, 16:04:59 : Search time 27.1004 seconds

(without alignments)
553,312 Million cell updates/sec

File: US-09-834-309-6

Perfect score: 752

Sequence: 1 EISDTPPEVKKARPPYSL.....ANEMWGPALPWCSEDFPLE 135

Scoring table: BLOSUM62

Gapop 10.0, Gapext 0.5

Searched: 747574 seqs, 1110706 residues

Total number of hits satisfying chosen parameters: 747574

Minimum db seq length: 0

Maximum db seq length: 2000000000

Post processing: Minimum Match 0%

Maximum Match 100%

Listing first 100 summaries

Database :				A_GenSeq 032802.*			
Result No.	Score	Query Match	Length	DB ID	Description	Pred. No.	Score
1	466.5	62.0	1087	12	AAE11149	8	lymphocyte memb
2	425.5	30.0	581	12	AAE13490	Human C4 binding	
3	421.5	29.5	2049	14	AAE36743	CR1_Homo sapien	
4	420.5	29.4	1540	12	AAE11982	Partial human com	
5	420.5	29.4	1926	22	AAE30184	Novel human diagn	
6	420.5	29.4	1916	19	AAE45899	Human complement	
7	420.5	29.4	2019	12	AAE11810	Human complement	
8	420.5	29.4	2019	20	AAE55751	Human C3b/C4b rec	
9	420.5	29.4	2044	22	AAE300287	Novel human diagn	
10	420.5	29.4	2044	22	AAE11782	Human CR1 protein	
11	420.5	29.4	2044	22	AAE39224	Human polypeptide	

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

12	220.5	23.3	2044	22	AAE41010	Human polypeptide
13	220.5	23.3	2317	19	AAE92219	CR1 protein_Homo
14	215	23.6	579	19	AAE39224	Amino acid sequen
15	212.5	23.3	450	20	AAE55755	Human CR1 protein
16	204.5	27.2	646	20	AAE55758	Human CR1 protein
17	199.5	25.5	363	18	AAE12414	Porcine complement
18	199.5	25.5	363	20	AAE30918	CR1 protein_Huid
19	197.5	23.3	263	26	AAE20858	Vaccinia virus SPI
20	197.5	23.3	263	20	AAE29860	Mutated VCP giving
21	195.5	23.3	263	22	AAE48846	Complex virus infla
22	193.5	23.7	263	10	AAE92003	Induced sequence o
23	193.5	23.7	263	20	AAE29859	Vaccinia complement
24	193.5	23.7	263	21	AAE14014	Complement inhibit
25	189.5	23.2	324	12	AAE15232	CR14 from clone pm
26	189.5	23.2	336	12	AAE15232	CR14 from clone pm
27	189.5	23.2	350	12	AAE15231	CR14 from clone pm
28	189.5	23.2	357	12	AAE19927	Human membrane cof
29	189.5	23.2	373	12	AAE15230	CR14 from clone pm
30	189.5	23.2	377	12	AAE15229	CR14 from clone pm
31	188.5	23.1	251	14	AAE41361	Membrane cofactor
32	188.5	23.1	254	17	AAE06880	Membrane cofactor
33	188.5	23.1	279	14	AAE41360	Membrane cofactor
34	188.5	23.1	279	19	AAE69523	CR14 protein SEQ
35	188.5	23.1	476	17	AAE94939	CR14 wild type_H
36	188.5	23.1	377	17	AAE94942	CR14 construct sub
37	188.5	23.1	377	17	AAE94941	CR14 construct sub
38	188.5	23.1	377	18	AAE27484	Human MCP_Homo S
39	188.5	23.1	378	17	AAE94940	CR14 construct sub
40	188.5	23.1	384	12	AAE10924	Human membrane cof
41	188.5	23.1	384	16	AAE86316	Human CR14_Homo
42	188.5	23.1	421	21	AAE54294	Local cap seq_38801
43	188.5	23.1	421	22	AAE37528	Human CR14 construct
44	188.5	23.1	577	17	AAE06882	Membrane cofactor
45	188.5	23.1	611	22	AAE15669	CR14 protein_Hui
46	188.5	23.1	611	22	AAE03762	CR14 construct pro
47	187.5	23.6	314	17	AAE94943	CR14 construct del
48	185.5	24.7	645	21	AAE53125	CR14 construct rha
49	178.5	23.7	133	15	AAE47156	Sequence of solubl
50	178.5	23.7	254	15	AAE47154	Sequence of solubl
51	178.5	23.7	438	20	AAE55756	Human CR1 protein
52	178.5	23.7	450	20	AAE55753	Human CR1 protein
53	178.5	23.7	476	20	AAE55752	Human CR1 protein
54	178.5	23.7	778	19	AAE73147	Amino acid sequen
55	175.5	23.3	254	15	AAE47155	Sequence of solubl
56	175.5	23.3	481	13	AAE20992	CR1-4 (amino acids
57	175.5	23.3	483	13	AAE20991	CR1-4 (amino acids
58	175.5	23.3	543	13	AAE28543	CR1-4 (45E, 37Y) a
59	175.5	23.3	543	13	AAE28544	CR1-4 (45E) analog
60	175.5	23.3	543	13	AAE28545	CR1-4 (37Y) analog
61	175.5	23.3	543	13	AAE28546	CR1-4 (44T, 47T, 4
62	175.5	23.3	543	13	AAE28547	CR1-4 (52S, 54S, 5
63	175.5	23.3	543	13	AAE28548	CR1-4 (57V, 59K) a
64	175.5	23.3	543	13	AAE28549	CR1-4 (64K, 65I) a
65	175.5	23.3	543	13	AAE28550	CR1-4 (64K) analog
66	175.5	23.3	543	13	AAE28551	CR1-4 (68T) analog
67	175.5	23.3	543	13	AAE28552	CR1-4 (78T, 79T) a
68	175.5	23.3	543	13	AAE28553	CR1-4 (84R, 87N) a
69	175.5	23.3	543	13	AAE28554	CR1-4 (92I, 94H) a
70	175.5	23.3	543	13	AAE28555	CR1-4 (92I) analog
71	175.5	23.3	543	13	AAE28556	CR1-4 (94H) analog
72	175.5	23.3	543	13	AAE28557	CR1-4 (99H, 103E)
73	175.5	23.3	543	13	AAE28558	CR1-4 (109N, 113A,
74	175.5	23.3	543	13	AAE28559	CR1-4 (114, 117S)K
75	175.5	23.3	543	13	AAE28560	CR1-4 (114T) analog
76	175.5	23.3	543	13	AAE28561	CR1-4 (115I) analog
77	175.5	23.3	543	13	AAE28562	CR1-4 (116K) analog
78	175.5	23.3	543	13	AAE28563	CR1-4 (117I) analog
79	175.5	23.3	543	13	AAE28564	CR1-4 (116K, 117F)
80	175.5	23.3	543	13	AAE28565	CR1-4 (121C) analog
81	175.5	23.3	543	13	AAE28566	CR1-4 (118R, 119N)
82	175.5	23.3	543	13	AAE28567	CR1-4 (11H, 12I) R
83	175.5	23.3	543	13	AAE28568	CR1-4 (147I, 149Y)
84	175.5	23.3	543	13	AAE28569	CR1-4 (169, 176) STK

86 106 5 20.3 543 1 AAR28570
 87 106 5 20.3 543 1 AAR28571
 88 169 22.5 667 15 AAR50087
 89 168 5 22.4 450 20 AAY55754
 90 168 5 22.4 453 20 AAY55757
 91 164 5 21.9 1655 22 AUC11860
 92 163 5 21.7 764 20 AAW94376
 93 162 5 21.6 600 22 AAI61858
 94 162 5 21.6 764 20 AAW94380
 95 162 21.5 920 22 AAG55921
 96 161 21.4 792 21 AAI42225
 97 161 21.4 982 22 AIG31221
 98 161 21.4 990 22 AIG55922
 99 161 21.4 990 22 AIC22405
 100 161 21.4 996 22 AIC55925

ALIGNMENTS

RESULT 1
 AAR11189
 ID AAR11189 standard; Protein 1087 AA.

AC AAR11189;

XX 23 MAY 1991 (first entry)

XX B lymphocyte membrane glycoprotein CR2.

DE CR2, B lymphocyte membrane receptor protein, Epstein Barr virus,
 KW extracellular domain.

XX Homo sapiens

XX Key Location/Qualifiers

FT Region 20..90

FT Znotc "1"

FT Znotc "2"

FT Znotc "3"

FT Znotc "4"

FT Znotc "5"

FT Znotc "6"

FT Znotc "7"

FT Znotc "8"

FT Znotc "9"

FT Znotc "10"

FT Znotc "11"

FT Znotc "12"

FT Znotc "13"

FT Znotc "14"

FT Znotc "15"

FT Znotc "16"

FT Znotc "17"

FT Znotc "18"

FT Znotc "19"

FT Znotc "20"

FT Znotc "21"

FT Znotc "22"

FT Znotc "23"

CR1 4 (266-274) KIK
 CR1 4 (364-467) NAA
 MCP-DAF fusion pro
 Human CR1 protein
 Human CR1 protein
 Prosopilia melanog
 Human Factor B and
 Myxoma virus immun
 Human Factor B and
 Human Factor B and
 Novel human diapo
 Novel human diapo
 Human OREF2489
 Novel human diapo
 Novel human diapo
 Novel human diapo
 Novel human diapo

PS Disclosure: Fig 9.1-9.9; 60pp; English.

XX The six indicated fragments encode claimed and disclosed

polypeptides, which are synthesized by recombinant expression, pref.

in a baculovirus expression system in which a DNA plasmid, contain

cDNA encoding this sequence, is truncated to encode a CR2 polypeptide

that comprises a region of the extracellular domain of CR2. The

construct is then inserted downstream of an appropriate promoter in

a transfect vector and integrated into a baculovirus which is then

used to infect host insect cells for expression.

XX The polypeptides correspond to B-lymphocyte membrane receptor

protein for Epstein-Barr Virus (EBV) and give specific binding.

XX Protein number 3 is used to inhibit infection of mammalian cells in

contact with an aq. medium, (esp. mammalian blood) and of human

B lymphocytes, by EBV. The peptides are used to detect the presence

of EBV in an aq. sample and to detect antibodies directed against CR2.

XX The peptides are used in a pharmacological compsh. as active

ingredient with a carrier to treat immune disorders.

XX Sequence 1087 AA;

Query Match 62.0%; Score 466.5; 108 12; Length 1087;

Best Local Similarity 61.2%; Pred. No. 6; 46 46;

Matches 82; Conservative 18; Mismatches 44; Indels 1; Gaps 1;

QY 2 ISGPPPPVNAPEYYSIPIVEHIVYIISPSYELINGKAKHICISFNAVNAWIRAPP 61

IR 21 ISGPPPPVNAPEYYSIPIVEHIVYIISPSYELINGKAKHICISFNAVNAWIRAPP 80

QY 62 ICESVNETISSEPIVTEHMESEFALFELFVHICISFNAVNAWIRAPP 121

IR 81 KCEFNKTSSEPIVTEHMESEFALFELFVHICISFNAVNAWIRAPP 130

QY 122 PIALVWESDEPLE 135

IR 140 PIRLPTVSVFELE 154

RESULT 2

AAR13490

XX AAR13490 standard; Protein: 581 AA.

AC AAR13490;

XX 30 OCT 1991 (first entry)

XX Human C4 binding protein.

XX C4bp, monomer; complement protein; pAb.C4bp.3; SCR5

XX short consensus repeat.

XX Homo sapiens.

XX Key Location/Qualifiers

FT Peptide 1..32

FT Peptide /label= signal_peptide

FT Protein 33..581

FT Region /label= C4bp

FT Region /label= SCR8

FT Region /label= SCR7

FT Region /label= SCR6

FT Region /label= SCR5

FT Region /label= SCR4

FT Region /label= SCR3

FT Region /label= SCR2

FT Region /label= SCR1

XX the invention relates to isolated polynucleotide (I) and
 CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
 CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
 CC and gene mapping, and in recombinant production of (II). The
 CC polynucleotides are also used in diagnostics as expressed sequence tags
 CC for identifying expressed genes. (I) is useful in gene therapy techniques
 CC to restore normal activity of (I) or to treat disease states involving
 CC (II). (II) is useful for generating antibodies against it, detecting or
 CC quantitating a polypeptide in tissue, as molecular weight markers and as
 CC a food supplement. (II) and its binding partners are useful in medical
 CC imaging of sites expressing (II). (I) and (II) are useful for treating
 CC disorders involving aberrant protein expression or biological activity.
 CC the polypeptide and polynucleotide sequences have applications in
 CC diagnostics, forensics, gene mapping, identification of mutations
 CC responsible for genetic disorders or other traits to assess biodiversity
 CC and to produce other types of data and products dependent on DNA and
 CC amino acid sequences. ABB00010 ABB00017 represent novel human
 CC diagnostic amino acid sequences of the invention.
 CC Note: The sequence data for this patent did not appear in the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at http://wipo.int/pub/published_pat_sequences.

XX Sequence 1929 AA;

Query Match 29.34; Score 220.5; DB 22; Length 1929;
 Best Local Similarity 36.04; Pred. No. 5,3e-13;
 Matches 49; Conservative 18; Mismatches 54; Indels 15; Gaps 5;

QY 2 ISGDPPEVKNAPKPYYS--LPIVPCIVLYICSPS-----YELI-EKATIGISENLY 52
 DB 1407 ISGDPPTISNG--DFTSNNTISPHNGIVYVYQGHGUGHQIFEDVWGRSIVGSKDDQ 1404
 QY 53 HATWKRAPPICISVNKTISSGPTIVPGRFMNKGSKAPFPRIDSVTFGRKANFTMGSKTV 112
 DB 1465 VGVWSSPPPRCISTNK--CTAPEVENATRYPNGRSPFSFTEIIFRCQPCFVWGSHTV 1521
 QY 113 WCOANEMKGPITALPVC 128
 DB 1522 QVQTRKRWGE KLPFC 1546

RESULT 6

AAW45899
 ID AAW45899 standard; peptide; 1930 AA;

XX AAW45899;

DE 30-JUN-1998 (first entry)

XX Human complement receptor 1 (residues 1-1929).

XX Membrane binding element; thrombotic disease; soluble protein;
 CC complement related disease; integral membrane protein; inflammation;
 CC short consensus repeat; SRK 1-3; CR1; complement receptor type 1.

OS Homo sapiens

XX Key Location/Qualifiers
 CC Cross-Links 1930
 CC /note= "Disulphide linked to Cys in peptide given
 CC in AAW45889"

XX W0902454 A2.

XX 22 JAN 1998.

XX 08 JUL 1997; 97WO-0004715.

XX 15 JUL 1996; 96GB-0014871.

XX (ADPR) ALPROTECH PLC.

XX

PI Todd I., Mossakowska DEL, Smith RAG;
 XX WPI, 1998-110524/10.

XX Derivatives of soluble poly:peptide(s) bonded to low affinity
 CC membrane binding groups - useful for treating complement related and
 CC thrombotic diseases, providing improved localisation at cellular
 CC membranes

XX Claim 22; Pages 60-61; 75pp; English.

XX This sequence represents human complement receptor 1 (CR1, Clb 45)
 CC N-terminal fragment. The invention relates to a soluble derivative (A)
 CC of a soluble polypeptide (I), which comprises at least 2 heterologous
 CC membrane-binding elements (MBE) of low membrane affinity covalently
 CC associated with (I). MBE interact, independently and with thermodynamic
 CC additivity, with components of cellular or artificial membranes exposed
 CC to extracellular fluids. (A) are used to treat disorders treatable with
 CC (I) itself, specifically inflammation or any other complement related
 CC disorder (e.g. thrombotic disease, graft rejection, myocardial
 CC infarction, sepsis, rheumatoid arthritis and many others) including
 CC application to indwelling devices) and thrombotic disease, but also to
 CC treat allergy induce weight loss, to treat ischaemia or asthma and as
 CC immuno-modulators for treating multiple sclerosis. (A) are administered
 CC orally, topically, by injection or inhalation at 0.01-10 (preferably
 CC 0.1-10) mg/kg/day.

XX Sequence 1930 AA;

Query Match 29.34; Score 220.5; DB 19; Length 1930;
 Best Local Similarity 36.04; Pred. No. 5,3e-13;
 Matches 49; Conservative 18; Mismatches 54; Indels 15; Gaps 5;

QY 2 ISGDPPEVKNAPKPYYS--LPIVPCIVLYICSPS--YELI-EKATIGISENLY 52
 DB 1476 ISGDPPTISNG--DFTSNNTISPHNGIVYVYQGHGUGHQIFEDVWGRSIVGSKDDQ 1534
 QY 53 HATWKRAPPICISVNKTISSGPTIVPGRFMNKGSKAPFPRIDSVTFGRKANFTMGSKTV 112
 DB 1534 VGVWSSPPPRCISTNK--CTAPEVENATRYPNGRSPFSFTEIIFRCQPCFVWGSHTV 1590
 QY 113 WCOANEMKGPITALPVC 128
 DB 1591 QCOITNIRWGI-KLPFC 1605

RESULT 7

AAW11810
 ID AAW11810 standard; Protein; 2039 AA

XX AAW11810;

XX 25-JUN-1991 (first entry)

XX Human complement type 1 receptor.

XX complement system; C3b/iC4b receptor; CR1; allergic reaction;
 CC immune response; clone lambda T109.1

OS Homo sapiens.

XX Key Location/Qualifiers
 CC Reptide 1..41
 CC /label= putative signal peptide
 CC Protein 42..2039
 CC /label= CR1

XX W09105047 A.

XX 19-APR-1991.

XX 25-SEP-1990; 90WO-US05454.

XX

KW C4B/C4B receptor; CRI protein; cell surface protein; erythrocyte; human;
KW complement regulatory activity; complement pathway enzyme; tissue damage;
KW reperfusion injury; Arthus reaction; myocardial infarct; inflammation;
KW heart condition; autoimmune disorder; long homologous repeat; LRP
XX
CS Homo sapiens.
XX
IN US981481 A.
XX
FD 09 NOV 1999.
XX
TF 06 JUN 1995; 95US-0470652.
PF
PF 05 APR 1989; 89US-0432865.
PF 06 DEC 1974; 74US-03450238.
PF 24 FEB 1994; 94US 0026134.
PF 01 APR 1988; 88US-0176532.
XX
PA (UYJO) UNIV JOHNS HOPKINS.
PA (HCHM) BRIGHAM & WOMEN'S HOSPITAL.
PA (AVAN) AVANT IMMUNOTHERAPEUTICS INC.
XX
PI Concilio MF, Wong WM, Makrides SC, Klickstein LB, Pearson DT, Ip SH;
PI Marsh RC, Carlson GR;
XX
DR WPI; 1999 63357/94.
XX
PT A human C4B/C4B receptor (CRI) protein having antiinflammatory and
PT cardiant activity.
XX
PS Disclosure; Fig 5b; 87pp; English.
XX
CC The invention relates to a human C4B/C4B receptor (CRI) protein. The CRI
CC protein or fragment is expressed as a cell-surface protein on the surface
CC of a non-human cell and exhibits a complement regulatory activity of full
CC length human CRI as expressed on erythrocytes. The CRI function in vivo
CC may be mediated through the inhibition of complement pathway enzymes. The
CC soluble CRI protein exhibits a complement regulatory activity, and this
CC may be used to prevent reperfusion injury, inhibit Arthus reaction, and
CC neutrophil mediated tissue damage, and reduce myocardial infarct size,
CC and inflammation. The CRI protein and its fragments can also be used in
CC the treatment of conditions which involve unwanted complement activity,
CC e.g. shock lung, tissue damage due to burn, or ischemic heart conditions,
CC and autoimmune disorders. CRI proteins, analogues, derivatives, and anti
CC CRI antibodies are used in assays, and diagnostics. The present sequence
CC represents the human CRI protein long homologous repeat (LHR)-D fragment.
XX
SQ Sequence 646 AA;
XX
Query Match 37.2%; Score 204.5; DB 20; Length 646;
Best Local Similarity 35.4%; Pred. No. 5.6e-12;
Matches 48; Conservative 16; Mismatches 57; Indels 15; Gaps 5;
QY 2 TSCDPPPEVKNARKPYYS---LPVDTVLRYTCSHS-----YRLIGKAIKFCISPNQV 52
DB 124 TSCDPPPEVKNARKPYYS---LPVDTVLRYTCSHS-----YRLIGKAIKFCISPNQV 181
QY 53 HATWDKAPPICESVYKNTISGSDPIVPGCFMKNKGSKAPFRHDSVTFCTCKANFTMKGSK 112
DB 182 VGVWSSPPPRCTISNK-CTAFVENATKVGNRRLFSLETFTRKRCQPCFVMVGSHTV 248
QY 113 WQANEMMGPTALPVG 128
DB 299 QQTNGRWGP-KLPQC 253
XX
RESULT 17
AAW12414
TO AAW12414 standard; Protein; 363 AA.
XX
AC AAW12414;
XX
DI 24 SEP 1997 (first entry)

XX Porcine complement inhibitor.
DE
XX porcine, pig; complement; inhibitor; organ transplantation;
KW analysis; promoter.
XX
OS Sus scrofa.
XX
PN WO9700951-A1.
XX
PD 09-JAN-1997.
XX
PF 19-JUN-1996; 96WO-JP01704.
XX
FR 20-JUN-1995; 95JP-0178254.
XX
PA (NIME-) NIPPON MEAT PACKERS INC.
PA (NIHA-) NIPPON HAW KK.
XX
PI Murakami H, Shigehisa T, Toyomura K;
XX
DR WPI; 1997-087378/08.
DR N-PSDE; AAT61098
XX
PT DNA encoding porcine complement inhibitor - useful in porcine organ
PT transplant to humans
XX
PS Claim 3; Page 2-14; 20pp; Japanese.
XX
CC This protein is a porcine complement inhibitor encoded by pmcPcDNA
CC (AAT61098). The DNA is useful for large scale production of
CC recombinant porcine complement inhibitor, which is useful for
CC porcine organ transplantation into humans. The DNA clone pmcPcDNA is
CC also useful in the analysis of the promoter region of porcine complement
CC inhibitor.
XX
SQ Sequence 363 AA;
XX
Query Match 26.5%; Score 199.5; DB 14; Length 363;
Best Local Similarity 31.9%; Pred. No. 8.4e-12;
Matches 45; Conservative 21; Mismatches 52; Indels 23; Gaps 6;
QY 1 ELSCHDPPPEVKNARK-----KPYYSLPVPGIVLRYTCSHS---YRLIGKAIKFCISPN 50
DB 168 FHLKPPGEPNGRYTNSHKTFEY-----NEVVTYSLSSTGDEFSLVGSSLEFISGR 223
QY 51 QVIATWDKAPPICESVYKNTISGSDPIVPGCFMKNKGSKAPFRHDSVTFCTCKANFTMKGSK 110
DB 224 E----WSSDPPPC---KVKKGPVVPVPCNCEIVSGP3SKFYVKAHVVPKCNACPFILHGRD 275
QY 111 TWWCQANEMMGPTALPVCESD 131
DB 276 TIVCGANSTWEE-EMPQCICK 295
XX
RESULT 18
AA330918
TO AAY40918 standard; Protein; 363 AA.
XX
AC AAY40918;
XX
DI 18-OCT-1999 (first entry)
XX
DE MCF protein.
XX
KW MCF, reporter gene; complement inhibitor; expression; antibody;
KW monoclonal antibody; IgG1; extracellular.
XX
OS Unidentified.
XX
PN JP11206392-A.
XX
DI 03-AUG-1999.

F1 Duplication 84
 F1 Znote "See note a in comments below."
 F1 Duplication 144
 F1 Znote "See note a in comments below."
 F1 Duplication 202
 F1 Znote "See note a in comments below."
 XX
 XX USN249208 N.
 XX
 XX 14 MAR 1989
 XX
 XX 20 AUG 1989 BRUS (249208)
 XX
 XX 20 AUG 1989 BRUS (249208)
 XX
 XX (USSH) NAT INST OF HEALTH
 XX
 XX Kozak
 XX
 XX WPI: 1989-06541/22
 XX N PSDB: AAN0113.
 XX
 XX New protein with anti-complement activity
 F1 encoded by Vaccinia virus 98K gene
 F1
 XX Disclosure: Figure 2A: copy: English.
 XX
 CC 64b binding protein which specifically blocks human complement cascades.
 CC It is the deduced sequence of a 64kDa protein encoded by sequence 52-840
 CC of the 98K gene of vaccinia virus strain WR. Note a - these sites
 CC indicate the start of 60 amino acid tandem repeating units which have a
 CC consensus sequence. The signal peptide sequence is not found in purified
 CC 98K protein recovered from the medium of cells infected with vaccinia
 CC virus strain WR. A suggested use is to treat diseases due to abnormally
 CC high complement activity.
 CC (Note: Revised entry submitted to correct the patent number format of
 CC US Government owned NIDB applications to prevent clashes with ongoing US
 CC granted patent numbers. For further information please visit the Derwent
 CC web site at www.derwent.com (45) relating to this as 0211)
 XX
 XX Sequence 263 AA:
 SQ
 Query Match 25.7% Score 193.5; DB 10; Length 263;
 Best local similarity 29.1% Pred. No. 2.4e-11;
 Matches 47; Conservative 19; Mismatches 60; Gaps 4;
 QY 2 ISCHPPEVKNAPKPYSLFVPGGLVETTSSTNRIICAKAFCTSHNVAHWKAPP 61
 DB 146 VKQSPSPSTSNCHNGYEDFYTHGSSVVTYSNSGYSLSGSCSG - GEWSO PP 199
 QY 62 PEEVKNKIFSCSGIVGGLMNSKATFRRGVSIFETANEIMKSKTWWQANEMW 121
 DB 200 IQV IVKCHDPTISNVTSSCHKRSYNGVDFKRYKELSSSSSPSPNTW 255
 QY 122 PIALPW 128
 DB 256 P-ELPKC 261
 QY 123
 AAY29859
 ID AAY29859 standard; protein: 263 AA.
 XX
 XX AAY29859;
 XX
 XX 16 NOV 1999 (first entry)
 XX
 XX Vaccinia complement control protein sequence.
 XX
 XX Vaccinia virus smallpox inhibitor of complement enzymes: SPIE;
 KW fusion protein: hyperacute rejection; xenograft; inflammation;
 KW post ischaemic reperfusion injury; malignancies; autoimmune disease;
 KW immune system disorder; neurodegeneration; infection; gene therapy;
 KW

KW blood additive; extracorporeal circulation system.
 XX
 XX Vaccinia virus.
 OS Synthetic.
 XX W09944625-A1.
 XX
 XX 10 SEP 1999.
 XX
 XX 02-MAR-1999; 9900 US04635.
 XX
 XX 03-MAR-1998; 9805-0076821.
 XX (OYPI) UNIV JOHNS HOPKINS.
 PA (OYPI) UNIV PITTSBURGH.
 XX
 XX Rosemond AM, Alcorn JM;
 XX WPI: 1999-050981/46.
 DR N-PSDB: AAZ21091.
 XX
 XX New smallpox inhibitor of complement enzyme proteins used to treat
 XX complement-mediated disease, particularly hyperacute rejection
 XX
 XX Claim 5; Fig 1; 88pp; English.
 XX
 CC The present invention describes the Vaccinia virus smallpox inhibitor of
 CC complement enzymes (SPIE) protein. SPIE is an inhibitor of complement
 CC activation, and so can be used to treat or prevent complement mediated
 CC disorders, especially hyperacute rejection, inflammation or post-
 CC ischaemic reperfusion injury, malignancies, autoimmune diseases,
 CC immune system disorders, neurodegeneration and infections. Hyperacute
 CC rejection may also be prevented by treating the graft with SPIE before
 CC transplanting it or by using a xenograft that has been transfected to
 CC express SPIE from a gene therapy vector. SPIE is also useful as
 CC additive to blood, e.g. in an extracorporeal circulation system (coated
 CC on tubing) or in storage, also for studying complement activation.
 CC Transgenic animals that express SPIE are used as sources of xenografts.
 CC The present sequence represents a vaccinia complement control protein
 CC (VCP) encoded by the specifically claimed mutated VCP nucleotide
 CC sequence, having a silent T to A transversion at nucleotide position
 CC number 267.
 XX
 XX Sequence 263 AA:
 SQ
 Query Match 25.7% Score 193.5; DB 20; Length 263;
 Best local similarity 29.1% Pred. No. 2.4e-11;
 Matches 47; Conservative 19; Mismatches 60; Gaps 4;
 QY 2 ISCHPPEVKNAPKPYSLFVPGGLVETTSSTNRIICAKAFCTSHNVAHWKAPP 61
 DB 146 VKQSPSPSTSNCHNGYEDFYTHGSSVVTYSNSGYSLSGSCSG - GEWSO PP 199
 QY 62 ICHSVKNTSCSGIVGGLMNSKATFRRGVSIFETANEIMKSKTWWQANEMW 121
 DB 200 IQV IVKCHDPTISNVTSSCHKRSYNGVDFKRYKELSSSSSPSPNTW 255
 QY 122 PIALPW 128
 DB 256 P-ELPKC 261
 RESULT 24
 AAB14014
 ID AAB14014 standard; protein: 263 AA.
 XX
 XX AAB14014;
 XX
 XX 11-DEC-2000 (first entry)
 XX
 XX Complement inhibitory protein VCP amino acid sequence.
 DE
 XX Alzheimer's disease; Vaccinia virus; VCP; complement pathway inhibitor.
 KW


```

DB 125 KVICTPPPKIKCKKHTFSKVEFEYLDVTV---YSCDHPAGDHPDPSLIGSTIYIC-GDNS 180
QV 52 VIATWCKAPPIESVKNKLTISCDPIVPGGFMNKGSKAPPRHDSVTETCKANFTMKSGSKT 111
DB 181 V---WSRAAPE2---KVKKHPFPVVENCKQISGFGKKFYKATVMPECDKGFYLDGSDT 233
QV 112 VMCQANFMCPITALVCSDEP 128
DB 234 IVCDNSTWDP-PVPKCLVLP 254

RESULT 29
AAR10927
ID AAR10927 standard; Protein: 457 AA.
XX A*
XX A*
XX 09 MAY 1991 (first entry)
XX Human membrane cofactor protein isoform with CYT1 C-terminus.
XX Mammalian membrane cofactor protein; complement activity;
XX inflammation; autoimmune diseases; tissue injury.
XX Key
XX Location/Qualifiers
XX 1..251
XX /label= short consensus repeats (4)
XX 252..266
XX /label= serine-threonine rich region, STA
XX 267..281
XX /label= STB
XX 282..295
XX /label= STC
XX 296..309
XX /label= unknown region
XX 310..341
XX /label= hydrophobic region
XX 342..357
XX /label= CYT1 cytoplasmic region
XX
XX W09102002 A.
XX 21 FEB 1991.
XX 20 JUL 1990; 90WO-0504107.
XX 19 APR 1990; 90NS-0510709.
XX 21 JUL 1989; 89US-0484210.
XX (UNIV) UNIV OF WASHINGTON.
XX Atkinson JP;
XX WPI: 1991-074491/10.
XX
XX New recombinant mammalian membrane cofactor protein - for
XX treating diseases with altered complement activity e.g.
XX inflammatory and autoimmune conditions
XX
XX Disclosure; fig 8; 38pp; English.
XX
XX This is a human isoform of a membrane cofactor protein (MCP).
XX Pharmaceutical compns. contg. the protein are useful in the
XX treatment of inflammatory and autoimmune diseases mediated by
XX excess or misdirection of complement activity, e.g. rheumatoid
XX arthritis or multiple sclerosis. Protection against tissue
XX injury caused by e.g. myocardial infarction or stroke may also
XX be provided. Antibodies raised against the MCP can be used in
XX autoimmune disease diagnosis to predict the probability of
XX recurrent miscarriages by testing for MCP levels in the placenta.
XX See also AAK10864-66.
XX
XX Sequence 457 AA:

```

```

Query Match 25.2%; Score 189.5; DB 12; Length 457;
Best Local Similarity 30.2%; Pred. No. 8.7e-11; Indels 15; Gaps 6;
Matches 42; Conservative 23; Mismatches 59;
QV 1 FISCDDPPPKVKNARKHYVSLPIVPG-TVIRYTCSPS-----YRLIGKKAIFCISENOVHA 54
DB 125 KVICTPPPKIKCKKHTFSKVEFEYLDVTVSYCDHPAGDHPDPSLIGSTIYIC-GDNSV-- 181
QV 55 TWDKAPPIESVKNKLTISCDPIVPGGFMNKGSKAPPRHDSVTETCKANFTMKSGSKTVMC 114
DB 182 -WSRAAPE2---KVKKHPFPVVENCKQISGFGKKFYKATVMPECDKGFYLDGSDTIVC 236
QV 115 QANFMCPITALVCSDEP 133
DB 237 DSNSTWDP-PVPKCLVLP 254

RESULT 29
AAR15230
ID AAR15230 standard; Protein: 373 AA.
XX XX
XX AAR15230;
XX 17-MAR-1992 (first entry)
XX CD46 from clone pm5.6.
XX
XX SGR: short consensus repeats; transmembrane; cytoplasmic;
XX membrane cofactor protein; MCP.
XX Homo sapiens.
XX Key
XX Location/Qualifiers
XX 1..34
XX /label= sig_peptide
XX 34..373
XX /label= mat_protein
XX 83
XX Modified-site
XX /label= N-glycosylation_site
XX 114
XX Modified-site
XX /label= N-glycosylation_site
XX 273
XX Modified-site
XX /label= N-glycosylation_site
XX 287..314
XX /note= "Ser/Thr rich region - O-linked
XX glycosylation sites"
XX 329..339
XX /label= hydrophobic_transmembrane_domain
XX
XX W09118097-A
XX
XX 28-NOV-1991
XX
XX 10-MAY-1991; 91WO-AU00199.
XX
XX 11-MAY-1990; 90AU-0000133.
XX (UYNF) UNIV MELBOURNE.
XX
XX Purcell DPJ, Russell SM, McKenzie JFC,
XX WPI: 1991-363251/50.
XX N-PSDB; AAK14916.
XX
XX New CD46 membrane co-factor protein variants - useful as probes
XX to identify CD46 isoforms and for diagnosing spontaneous
XX abortion, inhibiting immunoresponse and treating leukaemia
XX
XX Disclosure; Fig 1B and 3A; 77pp; English.
XX
XX The pm5.6 coding sequence corresponds to pm5.1, with the exception
XX that as for pm5.3, a 93 bp fragment is deleted from the fourth region

```


The sequences (AAQ46163-64) are derived from natural MCP (see patent WO91/02002) and are lacking either the transmembrane or intracellular region. Sequence (AAQ46164) contains the first 753 bases of (AAQ46163) plus a stop codon. Both sequences contain 4 SCR regions of natural MCP.

Primers (AAQ46165-66) and (AAQ46169-70) are used for amplification of the DNAs encoding these proteins.

XX Sequence 251 AA;

Query Match 25.1%; Score 188.5; DB 14; Length 251;
Best Local Similarity 40.6%; Pred. No. 7.2e-11;
Matches 41; Conservative 2; Mismatches 55; Indels 15; Gaps 6;

QY 1 EISCDPPPEVKNARKPYYSIPVPG-TVLRYTCS-----YRLIGKKAIFCISENOVHA 54

DB 125 KVLCTPPPKIKNGKHTESEVFEYLDAVYSCDPAFCDDPESLIGSTIYC-GDSNV-- 181

QY 55 TWKAPPICSVNKTISCSDPIVPGCFMNGSKAPPRHGDVSTPTCKANTMKGSKTVMC 114

DB 182 WSRAPPEC---KVKCRFPVVENGKQISCFCKKPYKATVMPFCDKGYLGGSDPTVC 236

QY 115 QANEMMGPTALPVC 128

DB 237 DSNSTWDP-PVPC 249

RESULT 42

AAQ46164

ID AAQ46164 standard; Protein: 254 AA.

XX AAQ46164;

XX AAQ46164;

DB 18 MAR-1997 (first entry)

DE Membrane co factor protein.

XX Complement inhibitor; membrane co-factor protein; MCP;

XX decay accelerating factor; DAF; chimeric protein; glycosaminoglycan;

XX heparin; cell lysis, sepsis; adult respiratory distress syndrome;

XX reperfusion injury; cell damage.

XX Homo sapiens.

XX W09644965-A2.

PN 07 NOV 1996.

XX 03 MAY 1996; 96WO 0506401.

XX 05 MAY 1995; 95US-0435149.

XX (CHIR) CHIRON CORP.

XX Creasey AA, Innis MA, Zaror T;

XX WPI; 1996 506167/40.

XX Chimeric proteins for inhibiting complement-mediated cell lysis -

XX comprise membrane co-factor protein and decay accelerating factor

XX peptide sequences

XX Disclosure; Page 25; 4pp; English.

XX A portion (AAQ46164) of the complement-inhibitor, membrane co-factor

XX protein (MCP), is used in novel chimeric proteins of formula

XX A-R1-B-R2-C, where A and C are peptides (AAQ46165-79, AAQ46169-90) able

XX to bind glycosaminoglycans (esp. heparin) present on cell surfaces,

XX R1 is a portion of MCP or decay accelerating factor (DAF, see also

XX AAQ46164), R2 is DAF when R1 is MCP or MCP when R1 is DAF, and B is a

XX peptide that may have complement inhibitor activity. The chimeric

XX proteins (see also AAQ46164) are directed to cell surfaces where they

XX inhibit complement mediated cell lysis. They are used to treat and

CC prevent disease states in which complement plays a role, e.g.

CC sepsis, adult respiratory distress syndrome, reperfusion injury and

CC tissue damage.

XX Sequence 254 AA;

Query Match 25.1%; Score 188.5; DB 17; Length 254;

Best Local Similarity 30.6%; Pred. No. 7.3e-11;

Matches 41; Conservative 23; Mismatches 55; Indels 15; Gaps 6;

QY 1 EISCDPPPEVKNARKPYYSIPVPG-TVLRYTCS-----YRLIGKKAIFCISENOVHA 54

DB 125 KVLCTPPPKIKNGKHTESEVFEYLDAVYSCDPAFCDDPESLIGSTIYC-GDSNV 181

QY 55 TWKAPPICSVNKTISCSDPIVPGCFMNGSKAPPRHGDVSTPTCKANTMKGSKTVMC 114

DB 182 WSRAPPEC---KVKCRFPVVENGKQISCFCKKPYKATVMPFCDKGYLGGSDPTVC 236

QY 115 QANEMMGPTALPVC 128

DB 237 DSNSTWDP-PVPC 249

RESULT 33

AAQ41360

ID AAQ41360 standard; Protein: 279 AA.

XX AAQ41360;

XX AAQ41360;

DT 09-MAR-1994 (first entry)

DE Membrane co factor protein fragment.

XX MCP; membrane co factor protein; transmembrane;

XX vascular inflammation; shock; graft rejection; SCR;

XX intracellular.

XX Homo sapiens.

XX W09317122-A.

PN 02-SEP-1993.

XX 22-FEB-1993; 93WO-3P00207.

XX 21-FEB-1992; 92JP-0035445.

XX (MOCH) MOCHIDA PHARM CO LTD.

XX Kanamori T, Nii A, Seya T;

XX WPI; 1993-288435/36.

XX N-PSDB; AAQ46163.

XX New protein comprising natural membrane co factor protein lacking

XX trans-membrane or intracellular region - or vascular

XX inflammation, shock, graft rejection, etc.

XX Claim 1; Page 15, 81pp, Japanese.

XX The sequences (AAQ46163-64) are derived from natural MCP

XX (see patent WO91/02002) and are lacking either the transmembrane or

XX intracellular region. Sequence (AAQ46164) contains the first 753 bases

XX of (AAQ46163) plus a stop codon. Both sequences contain 4 SCR regions

XX of natural MCP

XX Primers (AAQ46165-66) and (AAQ46169-70) are used for amplification of

XX the DNAs encoding these proteins.

XX Sequence 279 AA;

Query Match 25.1%; Score 188.5; DB 14; Length 279;

Best Local Similarity 30.6%; Pred. No. 7.3e-11;

Matches 41; Conservative 23; Mismatches 55; Indels 15; Gaps 6;


```

Db 215 WSKAAAFEC KVKRFPVVVNGKQISGFCCKKYYRAIVMFACUKGPFYLLGSDITVQ 269
QY 115 QANEMWGPTALPVC 128
Db 270 DSNSTWDP FVPRC 282

RESULT 36
AAR93942
ID AAR93942 standard; Protein: 377 AA.
XX AC
XX AAR93942;
XX DT 21-MAY-1996 (first entry)
XX DE CD46 construct subSCR2+3.
XX KW CD46; recombinant protein; short consensus repeat; SCR;
XX KW regulator of complement activation; transgenic animal; pig;
XX KW organ transplantation.
XX OS Synthetic.
XX PN WO9606937-A1.
XX PD 07-MAR-1996.
XX PF 30-AUG-1995; 95WO-AU00553.
XX PR 30-AUG-1994; 94AU-0007724.
XX PA (AUST-) AUSTIN RES INST.
XX PI Christiansen D, Loveland B, McKenzie IF, Milland J;
XX DR WO96/160368/16.
XX DR N-PSDB; AAT17598.
XX PT Increasing prodn. of recombinant proteins, esp. CD46 - by reducing
XX PT the amt. of A and/or T in an A and/or T rich region of encoding gene
XX PT exon
XX PS Claim 12; Page 48-49; 60pp; English.
XX CC CD46 subSCR2+3 (AAR93942) is the product of a cDNA construct
XX CC (AAT17598) obt'd. by splice overlap extension PCR of wild-type CD46
XX CC cDNA (AAT17595). The A/T contents of A-rich exons 3-5 of the
XX CC gene, encoding the short consensus repeats (SCR) 2 and 3 of
XX CC CD46, were lowered. This resulted in improved prodn. in eukaryotic
XX CC host cells, e.g. CHO-K1 and Cos-7. CD46 subSCR2+3 is used to
XX CC prevent complement or inflammation-mediated tissue damage, to
XX CC improve immunity to tumours or viruses, to control fertilisation
XX CC and to prevent spontaneous abortion. Expression in transgenic
XX CC animals, esp. pigs, provides organs suitable for transplantation.
XX SQ Sequence 377 AA;
Query Match 25.1%; Score 188.5; DB 17; Length 377;
Host Local Similarity 30.6%; Prod. No. 1.2e-10;
Matches 41; Conservative 23; Mismatches 55; Indels 15; Gaps 6;
QY 1 EISCDPPPEVKNAKPPYSLPIVPG-TLVRYTSPS-----YELIGEKATFCISENQVHA 54
Db 159 KVLCTPPPKIKNGKHITSEVEVEFYLDVATYSCDPAI(PDPFSLIGESTIYV-GDNSV 215
QY 55 TWKAKAPICFSVNRKISCSPPPIVGGCFMKNKGSKAFPHGLSVIFIPKAFNFMKSKIVWC 114
Db 216 WSKAAAFEC KVKRFPVVVNGKQISGFCCKKYYRAIVMFACUKGPFYLLGSDITVQ 270
QY 115 QANEMWGPTALPVC 128
Db 271 DSNSTWDP FVPRC 283

RESULT 36
AAR93942
ID AAR93942 standard; Protein: 377 AA.
XX AC
XX AAR93942;
XX DT 21-MAY-1996 (first entry)
XX DE CD46 construct subSCR2+3.
XX KW CD46; recombinant protein; short consensus repeat; SCR;
XX KW regulator of complement activation; transgenic animal; pig;
XX KW organ transplantation.
XX OS Synthetic.
XX PN WO9606937-A1.
XX PD 07-MAR-1996.
XX PF 30-AUG-1995; 95WO-AU00553.
XX PR 30-AUG-1994; 94AU-0007724.
XX PA (AUST-) AUSTIN RES INST.
XX PI Christiansen D, Loveland B, McKenzie IF, Milland J;
XX DR WO96/160368/16.
XX DR N-PSDB; AAT17598.
XX PT Increasing prodn. of recombinant proteins, esp. CD46 - by reducing
XX PT the amt. of A and/or T in an A and/or T rich region of encoding gene
XX PT exon
XX PS Claim 12; Page 48-49; 60pp; English.
XX CC CD46 subSCR2+3 (AAR93942) is the product of a cDNA construct
XX CC (AAT17598) obt'd. by splice overlap extension PCR of wild-type CD46
XX CC cDNA (AAT17595). The A/T contents of A-rich exons 3-5 of the
XX CC gene, encoding the short consensus repeats (SCR) 2 and 3 of
XX CC CD46, were lowered. This resulted in improved prodn. in eukaryotic
XX CC host cells, e.g. CHO-K1 and Cos-7. CD46 subSCR2+3 is used to
XX CC prevent complement or inflammation-mediated tissue damage, to
XX CC improve immunity to tumours or viruses, to control fertilisation
XX CC and to prevent spontaneous abortion. Expression in transgenic
XX CC animals, esp. pigs, provides organs suitable for transplantation.
XX SQ Sequence 377 AA;
Query Match 25.1%; Score 188.5; DB 17; Length 377;
Host Local Similarity 30.6%; Prod. No. 1.2e-10;
Matches 41; Conservative 23; Mismatches 55; Indels 15; Gaps 6;
QY 1 EISCDPPPEVKNAKPPYSLPIVPG-TLVRYTSPS-----YELIGEKATFCISENQVHA 54
Db 159 KVLCTPPPKIKNGKHITSEVEVEFYLDVATYSCDPAI(PDPFSLIGESTIYV-GDNSV 215
QY 55 TWKAKAPICFSVNRKISCSPPPIVGGCFMKNKGSKAFPHGLSVIFIPKAFNFMKSKIVWC 114
Db 216 WSKAAAFEC KVKRFPVVVNGKQISGFCCKKYYRAIVMFACUKGPFYLLGSDITVQ 270
QY 115 QANEMWGPTALPVC 128
Db 271 DSNSTWDP FVPRC 283

```


	ATOM	3737	C	LEU	C	45	62.664	85.961	-4.263	1.00	28.71	C
	ATOM	3738	O	LEU	C	45	63.094	86.894	-3.558	1.00	21.68	C
	ATOM	3739	N	CYS	C	46	61.410	85.547	-4.238	1.00	25.91	C
	ATOM	3740	CA	CYS	C	46	60.455	86.152	-3.343	1.00	26.90	C
5	ATOM	3741	C	CYS	C	46	60.610	85.392	-2.032	1.00	30.25	C
	ATOM	3742	O	CYS	C	46	60.448	84.173	-1.980	1.00	28.51	C
	ATOM	3743	CB	CYS	C	46	59.021	86.026	-3.908	1.00	28.02	C
	ATOM	3744	SG	CYS	C	46	57.741	86.546	-2.696	1.00	28.63	C
	ATOM	3745	N	ILE	C	47	60.925	85.115	-0.971	1.00	21.11	C
10	ATOM	3746	CA	ILE	C	47	61.120	85.484	0.314	1.00	24.65	C
	ATOM	3747	CD	ILE	C	47	60.649	85.581	-0.722	1.00	22.62	C
	ATOM	3748	CG2	ILE	C	47	63.505	84.878	-0.358	1.00	27.06	C
	ATOM	3749	CG1	ILE	C	47	63.050	85.958	0.959	1.00	25.64	C
	ATOM	3750	CD1	ILE	C	47	64.429	87.119	1.592	1.00	32.03	C
15	ATOM	3751	C	ILE	C	47	60.431	86.274	1.421	1.00	31.05	C
	ATOM	3752	O	ILE	C	47	59.931	87.399	1.226	1.00	22.58	C
	ATOM	3753	N	THR	C	48	60.421	85.669	2.596	1.00	27.58	C
	ATOM	3754	CA	THR	C	48	59.964	86.388	3.764	1.00	27.51	C
	ATOM	3755	CB	THR	C	48	58.584	85.987	4.273	1.00	28.85	C
	ATOM	3756	CG1	THR	C	48	59.313	86.770	5.120	1.00	25.12	C
	ATOM	3757	CG2	THR	C	48	58.490	84.526	4.619	1.00	26.50	C
	ATOM	3758	C	THR	C	48	61.001	86.164	4.839	1.00	29.22	C
	ATOM	3759	O	THR	C	48	61.455	85.745	5.045	1.00	24.24	C
	ATOM	3760	N	LYS	C	49	61.445	87.243	5.463	1.00	27.29	C
25	ATOM	3761	CA	LYS	C	49	62.425	87.112	6.555	1.00	31.22	C
	ATOM	3762	CB	LYS	C	49	63.478	88.220	6.427	1.00	26.53	C
	ATOM	3763	CG	LYS	C	49	64.229	88.239	5.134	1.00	33.44	C
	ATOM	3764	CD	LYS	C	49	65.703	88.351	5.346	1.00	40.12	C
	ATOM	3765	CE	LYS	C	49	66.262	87.134	6.055	1.00	42.32	C
30	ATOM	3766	NZ	LYS	C	49	67.729	87.262	6.166	1.00	46.64	C
	ATOM	3767	C	LYS	C	49	61.745	87.219	7.877	1.00	33.53	C
	ATOM	3768	O	LYS	C	49	62.249	86.692	8.856	1.00	30.36	C
	ATOM	3769	N	ASP	C	50	60.596	87.896	7.952	1.00	25.34	C
	ATOM	3770	CA	ASP	C	50	59.978	88.083	9.253	1.00	26.98	C
35	ATOM	3771	CB	ASP	C	50	59.961	89.579	9.597	1.00	25.16	C
	ATOM	3772	CG	ASP	C	50	59.154	90.397	8.607	1.00	27.41	C
	ATOM	3773	OD1	ASP	C	50	58.459	89.768	7.749	1.00	22.49	C
	ATOM	3774	OD2	ASP	C	50	59.207	91.664	8.693	1.00	26.37	C
	ATOM	3775	C	ASP	C	50	58.598	87.492	9.365	1.00	25.28	C
40	ATOM	3776	O	ASP	C	50	57.883	87.770	10.321	1.00	24.69	C
	ATOM	3777	N	LYS	C	51	58.247	86.664	8.386	1.00	26.19	C
	ATOM	3778	CA	LYS	C	51	56.948	86.006	8.319	1.00	23.97	C
	ATOM	3779	CB	LYS	C	51	56.717	85.091	9.526	1.00	31.09	C
	ATOM	3780	CG	LYS	C	51	57.749	84.021	9.749	1.00	32.84	C
45	ATOM	3781	CD	LYS	C	51	57.625	82.919	8.769	1.00	44.41	C
	ATOM	3782	CE	LYS	C	51	58.505	81.752	9.179	1.00	52.55	C
	ATOM	3783	NZ	LYS	C	51	58.266	81.397	10.605	1.00	51.40	C
	ATOM	3784	C	LYS	C	51	55.791	86.996	8.241	1.00	25.78	C
	ATOM	3785	O	LYS	C	51	54.660	86.628	8.510	1.00	25.03	C
50	ATOM	3786	N	VAL	C	52	56.048	88.252	7.896	1.00	25.78	C
	ATOM	3787	CA	VAL	C	52	54.937	89.195	7.758	1.00	23.31	C
	ATOM	3788	CB	VAL	C	52	54.976	90.286	8.894	1.00	30.82	C
	ATOM	3789	CG1	VAL	C	52	53.786	91.197	8.802	1.00	27.39	C
	ATOM	3790	CG2	VAL	C	52	54.946	89.611	10.264	1.00	33.79	C
55	ATOM	3791	C	VAL	C	52	55.085	89.832	6.376	1.00	29.16	C

II AAW06882 standard; Protein: 577 AA.

XX AAW06882;

XX 18 MAR 1997 (first entry)

XX Membrane co-factor protein-decay accelerating factor hybrid.

XX Complement inhibitor; membrane co-factor protein; MCP;

KW decay accelerating factor; DAF; chimeric protein; glycosaminoglycan;

KW heparin; cell lysis; sepsis; adult respiratory distress syndrome;

KW reperfusion injury; cell damage.

XX Homo sapiens.

XX Key Location/Qualifiers

FT Region 1..254

F1 Label= MCP

F1 255..577

F1 Label= DAF

XX W09644965 A2.

XX 07 NOV 1996.

XX 03 MAY 1996; 96W0-US06401.

XX 05 MAY 1995; 95US-0445149.

XX (CHIR) CHIRON CORP.

XX Creasey AA, Innis MA, Zaror L.

XX WPI: 1996 596167/50.

XX N PSDB: AAT46065.

XX Chimeric proteins for inhibiting complement-mediated cell lysis -
comprise membrane co factor protein and decay accelerating factor
peptide sequences

XX Disclosure: Page 25-26; 34pp; English.

XX A hybrid protein (AAW06882) comprises portions (see also AAW06880-81)
of the complement-inhibitors membrane co-factor protein (MCP) and
decay accelerating factor (DAF). It can be used in novel chimeric
proteins also incorporating a peptide (AAW06885-79, AAW06883-90) able
to bind glycosaminoglycans (esp. heparin) present on cell surfaces.
The constructs are encoded by overlapping PCR (see also AAT46066-72)
using MCP DAF hybrid DNA (AAT46065) as template, and can be expressed
in e.g. insect cells. The chimeric proteins are directed to cell
surfaces where they inhibit complement-mediated cell lysis. They
are used to treat and prevent disease states in which complement
plays a role, e.g. sepsis, adult respiratory distress syndrome,
reperfusion injury and tissue damage.

XX Sequence 577 AA;

Query Match 25 18; Score 188.5; DB 17; Length 577;
Best Local Similarity 40 68; Pred. No. 2e-10;
Matches 41; Conservative 24; Mismatches 55; Indels 15; Gaps 6;

QY 1 EISDPPPEVNKARKPYVSLPIVPG-TVLYTCSPS-----YRLIGKKAIFCISNGVHA 54

125 KVLCTPPPKIKNGKHITSEVEVEFYLLAVTYSCDPAGPDPFSLIGSTIYC-GDNSV-- 181

55 TWKKAPEICESVNKRTICSDDPIVPGFMNKGSKAPFPRRGDSVFTCKANFTMGSKIVWC 114

182 WSRAPAEPT----KVKRPFPIVVENKQISGCKKKFYKAIWMECUKQFYLDSDPIYC 246

QY 115 QANEMWGDTALPVC 128

DB 247 DSNSTWDD-PVTRC 249

RESULT 45

AAE12569

ID AAE12569 standard; Protein: 611 AA.

XX AAE12569;

XX 03-JAN-2002 (first entry)

XX CAB2 protein.

XX Expression vector; crippled selectable marker; neomycin resistance;

KW H.V protein; human immunodeficiency virus; improved expression; CAB2;

KW CAB4; amplifiable selectable marker; dihydrofolate reductase; dhfr;

KW transgene; continuous cell line preparation

XX Unidentified.

OS US2001024807-A1.

XX 27-SEP-2001

XX 22-DEC-2000; 2000US-0748061.

XX (1-NOV-1999, 99US-162930P.

XX 30-DEC-1999, 99US-0475460.

XX (CHIR) CHIRON CORP.

XX Innis M, Scott RM;

XX WPI: 2001 638503/73.

XX N-PSDB: AAD20355.

XX New expression vector, useful for improving expression of transgene or
polypeptide, comprises 3 polynucleotides encoding crippled selectable
marker, heterologous polypeptide or second amplifiable selectable
marker -

XX Example 2, Fig 1; 27pp; English.

XX The invention relates to a new expression vector comprising a first
polynucleotide encoding a first, crippled selectable marker which
include sequences encoding antibiotic (neomycin) resistance containing
one or more crippling mutations; second polynucleotide encoding a
heterologous polypeptide of interest which is viral protein (e.g. an
HIV protein) or is CAB2 or CAB4; and a third polynucleotide encoding
a second amplifiable selectable marker (e.g. dihydrofolate reductase
dhfr). The expression vectors are useful for the efficient expression
of desired polypeptides or improving expression of a transgene of
interest. The transformed cells can be used in the preparation of
continuous cell lines in which the cells are essentially immortal or
for the preparation of established cell lines that have the potential
to be subcultured in vitro. The present sequence is CAB2 protein.
XX Aberrant splicing of CAB2 DNA was corrected by removing donor and
XX acceptor sites using overlapping PCR

XX Sequence 611 AA;

Query Match 25 18; Score 188.5; DB 22; Length 611;
Best Local Similarity 30 68; Pred. No. 2.1e-10;
Matches 41; Conservative 23; Mismatches 55; Indels 15; Gaps 6;

QY 1 EISDPPPEVNKARKPYVSLPIVPG-TVLYTCSPS-----YRLIGKKAIFCISNGVHA 54

159 KVLCTPPPKIKNGKHITSEVEVEFYLLAVTYSCDPAGPDPFSLIGSTIYC-GDNSV 215

55 TWKKAPEICESVNKRTICSDDPIVPGFMNKGSKAPFPRRGDSVFTCKANFTMGSKIVWC 114

216 WSRAPAEPT----KVKRPFPIVVENKQISGCKKKFYKAIWMECTKPYLIGSTIYC 270

QY 115 QANEMWGDTALPVC 128

09:03:45 USN:TWB:TWB:2002

Search completed: November 6, 2002, 10:06:41
140 items, 40 1004 pages

GenCore version 5.1.3
Copyright (c) 1994-2002 CompuGen Ltd.

OM protein protein search, using sw model

Run on: November 6, 2002, 16:04:59 : Search time 10.539 Seconds
(without alignments)

312.880 Million cell updates/sec

File: US-09-834-409-6

Protein score: 752

Sequence: 1 ETSCHPPEVKNAPHYYSI ANEMWGPALPVCSDFPLE 135

Scoring table: BLOSUM62

Gapop 10.0, Gapext 0.5

Searched: 241628 seqs, 24425594 residues

Total number of hits satisfying chosen parameters: 231628

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post processing: Minimum Match 0%

Maximum Match 100%

Listing first 100 summaries

Database : Issued_Patents_AA:*

- 1: /cqn2_6/prodata/1/iaa/5A-COMB.pep.*
- 2: /cqn2_6/prodata/1/iaa/5B-COMB.pep.*
- 3: /cqn2_6/prodata/1/iaa/6A-COMB.pep.*
- 4: /cqn2_6/prodata/1/iaa/6B-COMB.pep.*
- 5: /cqn2_6/prodata/1/iaa/6C-COMB.pep.*
- 6: /cqn2_6/prodata/1/iaa/backfiles.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	220.5	29.3	1466	6	Patent No. 5256642
2	220.5	29.3	1466	6	Patent No. 5472939
3	220.5	29.3	1537	6	Patent No. 5256642
4	220.5	29.3	1537	6	Patent No. 5472939
5	220.5	29.3	1847	6	Patent No. 5256642
6	220.5	29.3	1847	6	Patent No. 5472939
7	220.5	29.3	2039	6	Patent No. 5256642
8	220.5	29.3	2039	6	Patent No. 5472939
9	199.5	26.5	463	4	Sequence 2, Appl 1
10	192.5	25.6	263	4	Sequence 2, Appl 1
11	190.5	25.3	384	4	Sequence 2, Appl 1
12	188.5	25.1	169	1	Sequence 16, Appl 1
13	188.5	25.1	169	2	Sequence 16, Appl 1
14	188.5	25.1	254	1	Sequence 13, Appl 1
15	188.5	25.1	254	2	Sequence 13, Appl 1
16	188.5	25.1	254	2	Sequence 13, Appl 1
17	188.5	25.1	293	1	Sequence 16, Appl 1
18	188.5	25.1	293	2	Sequence 16, Appl 1
19	188.5	25.1	424	2	Sequence 46, Appl 1
20	188.5	25.1	470	2	Sequence 42, Appl 1
21	188.5	25.1	473	2	Sequence 44, Appl 1
22	188.5	25.1	479	2	Sequence 44, Appl 1
23	188.5	25.1	484	6	Sequence 2, Appl 1
24	188.5	25.1	577	2	Sequence 3, Appl 1
25	188.5	25.1	611	4	Sequence 32, Appl 1
26	178.5	23.7	134	2	Sequence 31, Appl 1
27	178.5	23.7	134	2	Sequence 31, Appl 1

28	178.5	23.7	254	2	US-08-356-61-29	Sequence 29, Appl 1
29	178.5	23.7	254	2	US-08-769-967A-29	Sequence 29, Appl 1
30	175.5	23.3	254	2	US-08-356-61-30	Sequence 30, Appl 1
31	175.5	23.3	254	2	US-08-769-967A-30	Sequence 30, Appl 1
32	163	21.7	86	6	5514582-41	Patent No. 5514582
33	153.5	20.4	124	6	5514582-38	Patent No. 5514582
34	153.5	20.4	265	2	US-08-177-109A-57	Sequence 57, Appl 1
35	153.5	20.4	265	2	US-08-687-706-57	Sequence 57, Appl 1
36	153.5	20.4	764	2	US-08-177-139A-2	Sequence 2, Appl 1
37	153.5	20.4	764	2	US-08-687-706-2	Sequence 2, Appl 1
38	150.5	20.0	126	6	5514582-35	Patent No. 5514582
39	145.5	19.3	126	6	5514582-43	Patent No. 5514582
40	144	19.1	197	2	US-08-356-61-27	Sequence 27, Appl 1
41	144	19.1	197	2	US-08-769-967A-27	Sequence 27, Appl 1
42	142	18.9	323	2	US-08-435-49-2	Sequence 2, Appl 1
43	142	18.9	324	1	US-08-310-16A-14	Sequence 14, Appl 1
44	142	18.9	324	1	US-08-888-71-14	Sequence 14, Appl 1
45	139	18.5	610	1	US-08-365-470-3	Sequence 3, Appl 1
46	139	18.5	610	3	US-08-209-668-19	Sequence 19, Appl 1
47	139	18.5	610	4	US-08-009-490A-84	Sequence 89, Appl 1
48	139	18.5	610	6	5217870-2	Patent No. 5217870
49	130.5	17.4	274	2	US-08-177-109A-58	Sequence 58, Appl 1
50	130.5	17.4	274	2	US-08-687-706-58	Sequence 58, Appl 1
51	127.5	17.0	1019	1	US-08-296-014A-4	Sequence 4, Appl 1
52	127.5	17.0	1019	2	US-08-596-405-4	Sequence 4, Appl 1
53	127.5	17.0	1083	1	US-08-877-620-4	Sequence 2, Appl 1
54	127.5	17.0	1083	1	US-08-296-014A-2	Sequence 2, Appl 1
55	127.5	17.0	1083	2	US-08-596-405-2	Sequence 2, Appl 1
56	127.5	17.0	1083	2	US-08-877-620-2	Sequence 2, Appl 1
57	126.5	16.8	120	6	5514582-36	Patent No. 5514582
58	125.5	16.7	196	3	US-08-824-692-32	Sequence 32, Appl 1
59	125.5	16.7	229	3	US-08-824-692-31	Sequence 31, Appl 1
60	125.5	16.7	290	3	US-08-824-692-29	Sequence 29, Appl 1
61	124	16.5	830	6	5378464-2	Patent No. 5378464
62	123.5	16.4	484	2	US-08-252-193-9	Sequence 9, Appl 1
63	123.5	16.4	484	3	US-09-276-197-9	Sequence 9, Appl 1
64	123.5	16.4	830	5	PCT-US91-05056-2	Sequence 2, Appl 1
65	120	16.0	207	2	US-08-640-977-5	Sequence 5, Appl 1
66	120	16.0	248	2	US-08-640-977-2	Sequence 2, Appl 1
67	120	16.0	248	2	US-08-640-977-4	Sequence 4, Appl 1
68	120	16.0	326	2	US-08-640-977-1	Sequence 1, Appl 1
69	119	15.8	830	1	US-08-110-158-4	Sequence 4, Appl 1
70	115.5	15.4	574	6	5378464-3	Patent No. 5378464
71	113	15.0	127	6	5514582-37	Patent No. 5514582
72	110.5	14.7	177	3	US-08-824-692-36	Sequence 36, Appl 1
73	110	14.6	181	2	US-08-640-977-3	Sequence 3, Appl 1
74	106	14.1	128	6	5514582-42	Patent No. 5514582
75	104	13.9	128	6	5514582-39	Patent No. 5514582
76	103	13.7	128	6	5514582-35	Patent No. 5514582
77	100.5	13.4	127	6	5514582-33	Patent No. 5514582
78	100.5	13.4	177	3	US-08-824-692-30	Sequence 30, Appl 1
79	98	13.0	62	1	US-08-210-267A-13	Sequence 13, Appl 1
80	98	13.0	62	1	US-08-688-675-13	Sequence 13, Appl 1
81	98	13.0	62	3	US-08-477-866-13	Sequence 13, Appl 1
82	96	12.8	240	3	US-08-824-692-23	Sequence 23, Appl 1
83	95	12.6	216	3	US-08-824-692-24	Sequence 24, Appl 1
84	94	12.5	84	4	US-08-793-418-6	Sequence 6, Appl 1
85	93.5	12.4	933	2	US-08-313-230-1	Sequence 1, Appl 1
86	93.5	12.4	933	5	PCT-US93-038-1	Sequence 1, Appl 1
87	93	12.4	62	1	US-08-210-267A-12	Sequence 12, Appl 1
88	93	12.4	62	1	US-08-688-675-12	Sequence 12, Appl 1
89	93	12.4	62	3	US-08-477-866-12	Sequence 12, Appl 1
90	91	12.0	76	2	US-08-356-61-28	Sequence 28, Appl 1
91	90.5	12.0	17	4	US-08-769-967A-28	Sequence 28, Appl 1
92	90	12.0	17	4	US-08-602-996A-31	Sequence 31, Appl 1
93	88.5	11.9	385	2	US-08-340-547-2	Sequence 2, Appl 1
94	88.5	11.9	385	2	US-08-461-543-2	Sequence 2, Appl 1
95	87.5	11.6	372	2	US-08-513-271-2	Sequence 2, Appl 1
96	87.5	11.6	372	6	5514582-2	Patent No. 5514582
97	84.5	11.2	123	3	US-08-824-692-37	Sequence 37, Appl 1
98	84.5	11.2	123	3	US-08-824-692-38	Sequence 38, Appl 1
99	83	11.0	145	2	US-08-640-977-6	Sequence 6, Appl 1
100	82.5	11.0	128	6	5514582-34	Patent No. 5514582

ALIGNMENTS

```

RESULT 1
5256642 5
?Patent No. 5256642
? APPLICANT: FEARON, DOUGLAS T.; KLIKSTEIN, LLOYD B.; WING,
WINNIE W.; CARSON, GERALD R.; CONNING, MICHAEL F.; LIP, STEPHEN
H.; MAKRIDES, SAVVAS; MARSH, HENRY C., JR.
? TITLE OF INVENTION: COMPOSITIONS OF SOLUBLE COMPLEMENT
RECEPTOR 1 (CR1) AND A THROMBOLYTIC AGENT, AND THE METHODS OF
USE THEREOF
? NUMBER OF SEQUENCES: 40
? CURRENT APPLICATION DATA:
? APPLICATION NUMBER: US/09/588,128
? FILING DATE: 24 SEP 1999
? PRIOR APPLICATION DATA:
? APPLICATION NUMBER: 412,745
? FILING DATE: 26 SEP 1989
? APPLICATION NUMBER: 332,865
? FILING DATE: 03 APR 1989
? APPLICATION NUMBER: 176,532
? FILING DATE: 01 APR 1988
? SEQ ID NO: 5
? LENGTH: 1466
? 5256642-5

Query Match: 29.9% Score: 220.5; DR: 6; Length: 1466
Best Local Similarity: 36.0% Pred. No. 1 Re 15;
Matches: 49; Conservative: 19; Mismatches: 54; Indels: 15; Gaps: 5;

QY 2 ISDPTPEVKNAPPYYS---LPIVPTVLTLYTTPS YRLGKAFPTISENV 52
DB 1015 ISCEPTTSSG DYSNNRSTSHNGIVVYQHFGDQGFELVGRSVYTSKDD 1072

QY 53 BATWKAPEPTESVNFILSSGATVPEEMKSKSAFPRKASVILQKAMPTMSKIV 112
DB 1073 VGVWSPPTPLSTNK CLAPVNAIRVPTCHRTTLLTTRCQITVWVGHIV 1129

QY 113 WQANMMGCPALVQ 128
DB 1140 QQTNGRWGP KLEPC 1144

RESULT 2
5472949 5
?Patent No. 5472949
? APPLICANT: FEARON, DOUGLAS T.; KLIKSTEIN, LLOYD B.; WING,
WINNIE W.; CARSON, GERALD R.; CONNING, MICHAEL F.; LIP, STEPHEN
H.; MAKRIDES, SAVVAS; MARSH, HENRY C., JR.
? TITLE OF INVENTION: METHOD OF TREATING COMPLEMENT
MEDIATED DISORDERS
? NUMBER OF SEQUENCES: 40
? CURRENT APPLICATION DATA:
? APPLICATION NUMBER: US/09/148,825
? FILING DATE: 19 OCT 1993
? PRIOR APPLICATION DATA:
? APPLICATION NUMBER: 588,128
? FILING DATE: 24 SEP 1990
? APPLICATION NUMBER: 412,745
? FILING DATE: 26 SEP 1989
? SEQ ID NO: 6
? LENGTH: 1166
? 5472949-5

Query Match: 29.9% Score: 220.5; DR: 6; Length: 1466
Best Local Similarity: 36.0% Pred. No. 1 Re 15;
Matches: 49; Conservative: 19; Mismatches: 54; Indels: 15; Gaps: 5;

```

```

Matches: 49; Conservative: 19; Mismatches: 54; Indels: 15; Gaps: 5;

QY 2 ISDPTPEVKNAPPYYS---LPIVPTVLTLYTTPS YRLGKAFPTISENV 52
DB 1015 ISCEPTTSSG DYSNNRSTSHNGIVVYQHFGDQGFELVGRSVYTSKDD 1072

QY 53 BATWKAPEPTESVNFILSSGATVPEEMKSKSAFPRKASVILQKAMPTMSKIV 112
DB 1073 VGVWSPPTPLSTNK CLAPVNAIRVPTCHRTTLLTTRCQITVWVGHIV 1129

QY 113 WQANMMGCPALVQ 128
DB 1140 QQTNGRWGP KLEPC 1144

RESULT 3
5256642 5
?Patent No. 5256642
? APPLICANT: FEARON, DOUGLAS T.; KLIKSTEIN, LLOYD B.; WING,
WINNIE W.; CARSON, GERALD R.; CONNING, MICHAEL F.; LIP, STEPHEN
H.; MAKRIDES, SAVVAS; MARSH, HENRY C., JR.
? TITLE OF INVENTION: COMPOSITIONS OF SOLUBLE COMPLEMENT
RECEPTOR 1 (CR1) AND A THROMBOLYTIC AGENT, AND THE METHODS OF
USE THEREOF
? NUMBER OF SEQUENCES: 40
? CURRENT APPLICATION DATA:
? APPLICATION NUMBER: US/09/588,128
? FILING DATE: 24 SEP 1999
? PRIOR APPLICATION DATA:
? APPLICATION NUMBER: 412,745
? FILING DATE: 26 SEP 1989
? APPLICATION NUMBER: 332,865
? FILING DATE: 03 APR 1989
? APPLICATION NUMBER: 176,532
? FILING DATE: 01 APR 1988
? SEQ ID NO: 5
? LENGTH: 1537
? 5256642-5

Query Match: 29.9% Score: 220.5; DR: 6; Length: 1537
Best Local Similarity: 36.0% Pred. No. 1 Re 15;
Matches: 49; Conservative: 19; Mismatches: 54; Indels: 15; Gaps: 5;

QY 2 ISDPTPEVKNAPPYYS---LPIVPTVLTLYTTPS YRLGKAFPTISENV 52
DB 1015 ISCEPTTSSG DYSNNRSTSHNGIVVYQHFGDQGFELVGRSVYTSKDD 1072

QY 53 BATWKAPEPTESVNFILSSGATVPEEMKSKSAFPRKASVILQKAMPTMSKIV 112
DB 1073 VGVWSPPTPLSTNK CLAPVNAIRVPTCHRTTLLTTRCQITVWVGHIV 1129

QY 113 WQANMMGCPALVQ 128
DB 1140 QQTNGRWGP KLEPC 1144

RESULT 4
5472949 5
?Patent No. 5472949
? APPLICANT: FEARON, DOUGLAS T.; KLIKSTEIN, LLOYD B.; WING,
WINNIE W.; CARSON, GERALD R.; CONNING, MICHAEL F.; LIP, STEPHEN
H.; MAKRIDES, SAVVAS; MARSH, HENRY C., JR.
? TITLE OF INVENTION: METHOD OF TREATING COMPLEMENT
MEDIATED DISORDERS
? NUMBER OF SEQUENCES: 40
? CURRENT APPLICATION DATA:
? APPLICATION NUMBER: US/09/148,825
? FILING DATE: 19 OCT 1993
? PRIOR APPLICATION DATA:
? APPLICATION NUMBER: 588,128
? FILING DATE: 24 SEP 1990
? APPLICATION NUMBER: 412,745
? FILING DATE: 26 SEP 1989

```

Query Match: 25.1%; Score 104.5; EB 2; Length 374;
Best Local Similarity: 30.6%; Pred. No. 9.7e-15;
Matches: 41; Conservative: 23; Mismatches: 5; Indels: 19;
1 FLSGDPDEVKNAKRYVYSPIVDGTVRVSSPS--SVRTGKAKLPSN

DB	159	KVLC	PPPKLNKRNKHSEVEVEFYLDATVTDPAFAPDFSLIGESTIY	GNLSV	215
QY	55	TKKADPTCSVNRKLTSCSHPIVGGCFMKKSGAEPFACSVETPCRNPTMKSKVW	114		
DB	216	WSGALAE-----KVNRKTPVVENKQISGSAKFTALVYFELGKGYFAGSDIIV	270		
QY	115	QANEMHWETALPVC	128		
DB	271	DSNSLWIPVPPKVC	283		

RESULT 22
US-08-528-057-2
; Sequence 2, Application US/08528057
Patent No. 5,445,715

2 APPLICANT: PURCELL, Damian F. J.
 2 APPLICANT: RUSSELL, Sarah M.
 2 APPLICANT: MCKENZIE, Ian F. C.
 2 TITLE OF INVENTION: CD46 VARIANTS
 2 NUMBER OF SEQUENCES: 46
 2 CORRESPONDENCE ADDRESS:
 2 ADDRESS: POLCY & Lardner
 2 STREET: 1000 X Street, N.W., Suite 500
 2 CITY: Washington
 2 STATE: D. C.
 2 COUNTRY: USA
 2 ZIP: 20007-5109
 2 COMPUTER READABLE FORM:
 2 MEDIUM TYPE: Floppy disk
 2 COMPUTER: IBM PC compatible


```

QY 50 NQVHATWKAPEFESVKNL ESQSDIVVGGPMNKGSKAPFHHQSDVTFCKANFIMKGS 109
  1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
DB 58 DUCVGVWSGFATQCTFNK CTFENVENGLIVSINRSLFSLNFWVEF6QDPEVMKGP 114
  1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
QY 110 KIVWQANEMWGPALPG 28
  1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
DB 179 RRVKQALNKKWEP-ELFSC 24
  1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1

RESULT 28
US 08 356 361 40
Sequence 29, Application US/08/056861
GENERAL INFORMATION:
APPLICANT: Smith, Richard A.G.
APPLICANT: Dodd, Ian
APPLICANT: Freeman, Mary A.
APPLICANT: Mossakowska, Danuta E.L.
TITLE OF INVENTION: Novel carbonyl compounds
NUMBER OF SEQUENCES: 41
CORRESPONDENCE ADDRESS:
ADDRESSEE: SmithKline Beecham Corporate Intellectual Property
STREET: P.O. Box 1549
CITY: Kind of Prussia
STATE: Pennsylvania
COUNTRY: USA
ZIP: 19406
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/056861
FILING DATE: 03 Jul 1995
CLASSIFICATION: 445
ATTORNEY/AGENT INFORMATION:
NAME: Robert H.
REGISTRATION NUMBER: 41141
REFERENCE/BOOKET NUMBER: P40424
TELEPHONE: (610) 270-5099
TELEFAX: (610) 270-5090
INFORMATION FOR SEQ ID NOS 29:
SEQUENCE CHARACTERISTICS:
LENGTH: 254 amino acids
TYPE: amino acid
TOPOLOGY: Linear
MOLECULE TYPE: Peptide
FRAGMENT TYPE: N-terminal
US 08 356 361 29

Query Match 24.7% Score 178.5; DB 2; Length 254;
Best Local Similarity 40.9%; Pred. No. 7; 46 12;
Matches 43; Conservative 20; Mismatches 55; Indels 21; Gaps 5;

QY 2 LSCDPEPVKNA RKFYSLPVLGVLPVYSPS YELDERALFISE 49
  1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
DB 124 LQGLPLTLINDEFISTENRHHY GSVVYFVNGSGGKRVFELGVGDSYICISN 178
  1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
QY 50 NQVHATWKAPEFESVKNL ESQSDIVVGGPMNKGSKAPFHHQSDVTFCKANFIMKGS 109
  1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
DB 179 DUCVGVWSGFATQCTFNK CTFENVENGLIVSINRSLFSLNFWVEF6QDPEVMKGP 24
  1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
QY 110 KIVWQANEMWGPALPG 28
  1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
DB 246 RRVKQALNKKWEP-ELFSC 254
  1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1

RESULT 29
US 08 356 361 29
Sequence 29, Application US/08/056861
GENERAL INFORMATION:
APPLICANT: Smith, Richard A.G.
APPLICANT: Dodd, Ian
APPLICANT: Freeman, Mary A.
APPLICANT: Mossakowska, Danuta E.L.
TITLE OF INVENTION: Novel carbonyl compounds
NUMBER OF SEQUENCES: 41
CORRESPONDENCE ADDRESS:
ADDRESSEE: SmithKline Beecham Corporate Intellectual Property
STREET: P.O. Box 1549
CITY: Kind of Prussia
STATE: Pennsylvania
COUNTRY: USA
ZIP: 19406
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/056861
FILING DATE: 03 Jul 1995
CLASSIFICATION: 445
ATTORNEY/AGENT INFORMATION:
NAME: Robert H.
REGISTRATION NUMBER: 41141
REFERENCE/BOOKET NUMBER: P40424
TELEPHONE: (610) 270-5099
TELEFAX: (610) 270-5090
INFORMATION FOR SEQ ID NOS 29:
SEQUENCE CHARACTERISTICS:
LENGTH: 254 amino acids
TYPE: amino acid
TOPOLOGY: Linear
MOLECULE TYPE: Peptide
FRAGMENT TYPE: N-terminal
US 08 356 361 29

```

```

Patent No. 5859223
GENERAL INFORMATION:
APPLICANT: Mossakowska, Danuta E.L.
APPLICANT: Smith, Richard A.G.
APPLICANT: Dodd, Ian
APPLICANT: Freeman, Anne Mary
TITLE OF INVENTION: Soluble CR1 derivatives
NUMBER OF SEQUENCES: 44
CORRESPONDENCE ADDRESS:
ADDRESSEE: SmithKline Beecham Corporate Intellectual Property
STREET: P.O. Box 1549
CITY: Kind of Prussia
STATE: Pennsylvania
COUNTRY: USA
ZIP: 19406
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/769,967A
FILING DATE:
CLASSIFICATION: 546
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/440,569
FILING DATE: 15-May-1995
ATTORNEY/AGENT INFORMATION:
NAME: K.04, William T.
REGISTRATION NUMBER: 30,954
REFERENCE/BOOKET NUMBER: P4042402
TELECOMMUNICATION INFORMATION:
TELEPHONE: (610) 270-5090
TELEFAX: (610) 270-5090
INFORMATION FOR SEQ ID NOS 29:
SEQUENCE CHARACTERISTICS:
LENGTH: 254 amino acids
TYPE: amino acid
TOPOLOGY: Linear
MOLECULE TYPE: peptide
FRAGMENT TYPE: N terminal
US 08 769 967A 29

Query Match 24.7% Score 178.5; DB 2; Length 254;
Best Local Similarity 40.9%; Pred. No. 7; 46 12;
Matches 43; Conservative 20; Mismatches 55; Indels 21; Gaps 5;

QY 2 LSCDPEPVKNA RKFYSLPVLGVLPVYSPS YELDERALFISE 49
  1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
DB 124 LQGLPLTLINDEFISTENRHHY GSVVYFVNGSGGKRVFELGVGDSYICISN 178
  1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
QY 50 NQVHATWKAPEFESVKNL ESQSDIVVGGPMNKGSKAPFHHQSDVTFCKANFIMKGS 109
  1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
DB 179 DUCVGVWSGFATQCTFNK CTFENVENGLIVSINRSLFSLNFWVEF6QDPEVMKGP 24
  1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
QY 110 KIVWQANEMWGPALPG 28
  1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
DB 246 RRVKQALNKKWEP-ELFSC 254
  1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1

RESULT 30
US 08 356 361 40
Sequence 30, Application US/08/056861
Patent No. 585989
GENERAL INFORMATION:
APPLICANT: Smith, Richard A.G.
APPLICANT: Dodd, Ian
APPLICANT: Freeman, Mary A.
APPLICANT: Mossakowska, Danuta E.L.
TITLE OF INVENTION: Novel carbonyl compounds
NUMBER OF SEQUENCES: 41
CORRESPONDENCE ADDRESS:
ADDRESSEE: SmithKline Beecham Corporate Intellectual Property

```


STREET: P.O. Box 1539
CITY: Kind of Prussia
STATE: Pennsylvania
COUNTRY: USA
ZIP: 19406
COMPUTER READABLE FORM:
MEDIUM TYPE: 10M floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/456,461
FILING DATE: 03-Jul-1995
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: JETVIS, Holbert H.
REGISTRATION NUMBER: 41,171
REFERENCE/DOCKET NUMBER: P40424
TELECOMMUNICATION INFORMATION:
TELEPHONE: (610) 270-5019
TELEFAX: (610) 270-5090
INFORMATION FOR SEQ ID NO: 30:
SEQUENCE CHARACTERISTICS:
LENGTH: 254 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
FRAGMENT TYPE: N-terminal
US 08 456-461-40

Query Match 23.4%; Score 175.5; DR 2; Length 254;
Best Local Similarity 30.9%; Pred. No. 1.5e-11;
Matches 43; Conservative 19; Mismatches 56; Indels 21; Gaps 5;

27 2 ISCDPPEVKNA-----RKPYSLPIVPGTIVRTYSPS-----VRLIGKAIPTISE 49
Db 124 LPGLPTTNGDFISTNNRPHY---GSSVVTYRNPSSGRKRVFFLVEEFSIYCTSN 178

QY 50 NOVHATWKAIPICSVNKTISCSDDPIVGGGNKRC:KAPFHGDSVFTCKANFTMKGS 109
Db 179 DQVGTWSGAPQCTIPNK---CTPPRVEN:ILVSEFSLF:NEVVEFCQGFVYMKGP 235

QY 110 KTVWCOANEMWGPTALPVC 128
Db 236 HRVKCOALNKWEP-ELPSC 253

RESULT 41
US 08 769-967A 40
Sequence 40, Application US/08769967A
Patent No. 5859224
GENERAL INFORMATION:
APPLICANT: Mossakowska, Danuta E.L.
APPLICANT: Smith, Richard A.G.
APPLICANT: Dodd, Ian
APPLICANT: Freeman, Anne Mary
TITLE OF INVENTION: Soluble "RI" Derivatives
NUMBER OF SEQUENCES: 44
CORRESPONDENCE ADDRESS:
ADDRESSEE: SmithKline Beecham Corporate Intellectual Property
STREET: P.O. Box 1539
CITY: Kind of Prussia
STATE: Pennsylvania
COUNTRY: USA
ZIP: 19406
COMPUTER READABLE FORM:
MEDIUM TYPE: 10M floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/769,967A
FILING DATE:

CLASSIFICATION: 536
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/440,569
FILING DATE: 15-May-1995
ATTORNEY/AGENT INFORMATION:
NAME: King, William T.
REGISTRATION NUMBER: 30,954
REFERENCE/DOCKET NUMBER: P40424C2
TELECOMMUNICATION INFORMATION:
TELEPHONE: (610) 270-5364
TELEFAX: (610) 270-5090
INFORMATION FOR SEQ ID NO: 30:
SEQUENCE CHARACTERISTICS:
LENGTH: 254 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
FRAGMENT TYPE: N-terminal
US-08-769-967A-30

Query Match 23.4%; Score 175.5; DR 2; Length 254;

Best Local Similarity 30.9%; Pred. No. 1.5e-11;
Matches 43; Conservative 19; Mismatches 56; Indels 21; Gaps 5;

QY 2 ISCDPPEVKNA-----RKPYSLPIVPGTIVRTYSPS-----VRLIGKAIPTISE 49

Db 124 LPGLPTTNGDFISTNNRPHY---GSSVVTYRNPSSGRKRVFFLVEEFSIYCTSN 178

QY 50 NOVHATWKAIPICSVNKTISCSDDPIVGGGNKRC:KAPFHGDSVFTCKANFTMKGS 109

Db 179 DQVGTWSGAPQCTIPNK---CTPPRVEN:ILVSEFSLF:NEVVEFCQGFVYMKGP 235

QY 110 KTVWCOANEMWGPTALPVC 128

Db 236 HRVKCOALNKWEP-ELPSC 253

RESULT 42

5514582-41

Patent No. 5514582

APPLICANT: CAPIN, DANIEL J.; LASKY, LAUREN E. A.

TITLE OF INVENTION: RECOMBINANT DNA ENCODING HYBRID

IMMUNOGLOBULINS

NUMBER OF SEQUENCES: 43

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/185,670

FILING DATE: 21-JAN-1994

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 986,941

FILING DATE: 08-DEC-1992

APPLICATION NUMBER: 808,122

FILING DATE: 16-DEC-1991

APPLICATION NUMBER: 440,625

FILING DATE: 22-NOV-1989

APPLICATION NUMBER: 315,015

FILING DATE: 23-FEB-1989

SEQ ID NO: 41

LENGTH: 86

5514582 41

Query Match 21.7%; Score 163; DR 6; Length 86;

Best Local Similarity 32.6%; Pred. No. 4.4e-11;
Matches 29; Conservative 15; Mismatches 41; Indels 4; Gaps 1;

QY 32 CSDPYRLIGKAIPTISENOVHATWKAIPICSVNKTISCSDDPIVGGGNKRC:KAPFHGDSVFTCKANFTMKGS 91

Db 1 CDPFSLGLGHASISCTVENETIGVWRPSPPT:EK---ITCRKPVSHSHVMSGFGPIYN 56

QY 92 HGSVFTCKANFTMKGSKTVWCOANEMW 120

Db 57 YKDTIVFKCQGFVIRCSSVIRHCDADSKW 85

```

RESULT 66
5014562 59
PATENT NO. 5014562
APPLICANT: CAPN, DANIEL J.; LASKY, LAURENCE A.
TITLE OF INVENTION: RECOMBINANT DNA ENCODING HYBRID
IMMUNOGLOBULINS
NUMBER OF SEQUENCES: 44
CURRENT APPLICATION DATA
  FILING DATE: 21 JAN 1994
  PUBLICATION NUMBER: 980,941
  FILING DATE: 08 DEC 1992
  APPLICATION NUMBER: 808,122
  FILING DATE: 16 DEC 1994
  APPLICATION NUMBER: 440,625
  FILING DATE: 22 NOV 1999
  APPLICATION NUMBER: 415,016
  FILING DATE: 23 FEB 1999
SEQ ID NO: 39
LENGTH: 124
Query Match
  Best Local Similarity 41.1% Score 15.4 59 DB 62 Length 124
  Matches 42: Conservative 17: Mismatches 59: Indels 17: Gaps 62
QY 2 ISCDPEPVNA PKYYSLPVIVYVYVSFSLGKAFGLFISNQVHATW 57
DB 1 LTRPDRHDEYEWKSPYVW SDFSHVYVGYTLRSANRIT ---QVNRKWS 57
QY 58 KAPFTEVNRKTSSTIVGEMKSGKAFGLFISNQVHATW 117
DB 54 QTAICD NAGYCSNPGIPGTRKVSQ YRLEDSVTVHSGRLTGRSRRPQEG 108
QY 118 EMMPDIALVWESDF 142
DB 109 GWSGATF ESQDF 122
Query Match
  Best Local Similarity 41.1% Score 15.4 59 DB 62 Length 124
  Matches 42: Conservative 17: Mismatches 59: Indels 17: Gaps 62
QY 2 ISCDPEPVNA PKYYSLPVIVYVYVSFSLGKAFGLFISNQVHATW 57
DB 1 LTRPDRHDEYEWKSPYVW SDFSHVYVGYTLRSANRIT ---QVNRKWS 57
QY 58 KAPFTEVNRKTSSTIVGEMKSGKAFGLFISNQVHATW 117
DB 54 QTAICD NAGYCSNPGIPGTRKVSQ YRLEDSVTVHSGRLTGRSRRPQEG 108
QY 118 EMMPDIALVWESDF 142
DB 109 GWSGATF ESQDF 122

```

```

RESULT 67
US 08 177-109A 57
Sequence 57, Application US/08/177/109A
Patent No. 5928492
GENERAL INFORMATION:
  APPLICANT: Jeanis E. Bourcade and Teresa J. Colosby
  TITLE OF INVENTION: MODIFIED COMPLEMENT PROTEASES
  NUMBER OF SEQUENCES: 62
  CORRESPONDENCE ADDRESS:
  ADDRESSEE: Patricia L. Pabst
  STREET: 2800 One Atlantic Center
  CITY: Atlanta
  STATE: Georgia
  COUNTRY: USA
  ZIP: 30309-3450
  COMPUTER READABLE FORM:
  MEDIUM TYPE: Floppy disk
  OPERATING SYSTEM: PC DOS/MS-DOS
  SOFTWARE: Patent in Release #1.0, Version #1.25
  CURRENT APPLICATION DATA:
  FILING DATE: 26 JUL 1996
  CLASSIFICATION: 514
  PRIOR APPLICATION DATA:
  APPLICATION NUMBER: US 08/177,109
  FILING DATE: 03-JAN-1994
  CLASSIFICATION: 514
  ATTORNEY/AGENT INFORMATION:
  NAME: Pabst, Patricia L.
  REGISTRATION NUMBER: 41,284
  REFERENCE/BOOKLET NUMBER: WI 107 D1V
  TELECOMMUNICATION INFORMATION:
  TELEPHONE: (404) 874-8794
  TELEFAX: (404) 874-8795
  INFORMATION FOR SEQ ID NO: 57:
  SEQUENCE CHARACTERISTICS:
  LENGTH: 265 amino acids
  TYPE: amino acid
  TOPOLOGY: linear
  MOLECULE TYPE: protein
  HYPOTHEetical: NO
  INFORMATION FOR SEQ ID NO: 57

```

```

SEQUENCE CHARACTERISTICS:
  LENGTH: 265 amino acids
  TYPE: amino acid
  TOPOLOGY: linear
  MOLECULE TYPE: protein
  HYPOTHEtical: NO
US-08 177-109A-57
Query Match
  Best Local Similarity 41.1% Score 15.4 59 DB 62 Length 265
  Matches 42: Conservative 17: Mismatches 59: Indels 17: Gaps 62
QY 2 ISCDPEPVNA PKYYSLPVIVYVYVSFSLGKAFGLFISNQVHATW 57
DB 1 LTRPDRHDEYEWKSPYVW SDFSHVYVGYTLRSANRIT ---QVNRKWS 127
QY 58 KAPFTEVNRKTSSTIVGEMKSGKAFGLFISNQVHATW 117
DB 128 QTAICD NAGYCSNPGIPGTRKVSQ YRLEDSVTVHSGRLTGRSRRPQEG 184
QY 118 EMMPDIALVWESDF 142
DB 184 GWSGATF ESQDF 197
RESULT 65
US 08 687-706 57
Sequence 57, Application US/08/687/06
Patent No. 5928492
GENERAL INFORMATION:
  APPLICANT: Jeanis E. Bourcade and Teresa J. Colosby
  TITLE OF INVENTION: MODIFIED COMPLEMENT PROTEASES
  NUMBER OF SEQUENCES: 62
  CORRESPONDENCE ADDRESS:
  ADDRESSEE: Patricia L. Pabst
  STREET: 2800 One Atlantic Center
  CITY: Atlanta
  STATE: Georgia
  COUNTRY: USA
  ZIP: 30309-3450
  COMPUTER READABLE FORM:
  MEDIUM TYPE: Floppy disk
  OPERATING SYSTEM: PC DOS/MS-DOS
  SOFTWARE: Patent in Release #1.0, Version #1.25
  CURRENT APPLICATION DATA:
  FILING DATE: 26 JUL 1996
  CLASSIFICATION: 514
  PRIOR APPLICATION DATA:
  APPLICATION NUMBER: US 08/177,109
  FILING DATE: 03-JAN-1994
  CLASSIFICATION: 514
  ATTORNEY/AGENT INFORMATION:
  NAME: Pabst, Patricia L.
  REGISTRATION NUMBER: 41,284
  REFERENCE/BOOKLET NUMBER: WI 107 D1V
  TELECOMMUNICATION INFORMATION:
  TELEPHONE: (404) 874-8794
  TELEFAX: (404) 874-8795
  INFORMATION FOR SEQ ID NO: 57:
  SEQUENCE CHARACTERISTICS:
  LENGTH: 265 amino acids
  TYPE: amino acid
  TOPOLOGY: linear
  MOLECULE TYPE: protein
  HYPOTHEtical: NO
  INFORMATION FOR SEQ ID NO: 57
US 08 687-706 57

```

```

Query Match
  Best Local Similarity 41.1% Score 15.4 59 DB 62 Length 265
  Matches 42: Conservative 17: Mismatches 59: Indels 17: Gaps 62

```


FILING DATE 08 DEC 1992
 APPLICATION NUMBER: 808 122
 FILING DATE 16 DEC 1991
 APPLICATION NUMBER: 430 626
 FILING DATE 22 NOV 1989
 APPLICATION NUMBER: 415 615
 FILING DATE 23 FEB 1989
 SEQ ID NO: 43
 LENGTH: 126
 G514562 5

Query Match
 Best Local Similarity 50.0% Score 150.5; DB 6; Length 126;
 Matches 40; Conservative 16; Mismatches 55; Indels 25; Gaps 8;

QY 4 CDEPPEVKNKPYSLFVPTSPSYELDEKALPQISENVIAWKRAFF 58
 DB 1 CDEPPEVKNKPYSLFVPTSPSYELDEKALPQISENVIAWKRAFF 58
 QY 54 ESNVKNLISNLSDFIVKGMNKSAPFRBDVSFTFCANETMKSRTVM 114
 DB 54 ESNVKNLISNLSDFIVKGMNKSAPFRBDVSFTFCANETMKSRTVM 114
 QY 115 CANEMMGTALAVESDF 126
 DB 115 CANEMMGTALAVESDF 126

RESULT 4
 G514562 43
 Patent No. G514562

APPLICANT: CAP N. DANIEL J. CLARKY LAURENCE A.
 TITLE OF INVENTION: RECOMBINANT DNA ENCODING HYBRID
 IMMUNOGLYCOPOLIN
 NUMBER OF SEQUENCES: 14
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: 05/007185, 670
 FILING DATE 24 JAN 1994

PREVIOUS APPLICATION DATA:
 APPLICATION NUMBER: 986,941
 FILING DATE 08 DEC 1992
 APPLICATION NUMBER: 808,122
 FILING DATE 16 DEC 1991
 APPLICATION NUMBER: 430,626
 FILING DATE 22 NOV 1989
 APPLICATION NUMBER: 415,615
 FILING DATE 23 FEB 1989
 SEQ ID NO: 43
 LENGTH: 126
 G514562 43

Query Match
 Best Local Similarity 26.2% Score 145.5; DB 6; Length 126;
 Matches 40; Conservative 22; Mismatches 62; Indels 9; Gaps 4;

QY 4 CDEPPEVKNKPYSLFVPTSPSYELDEKALPQISENVIAWKRAFF 64
 DB 4 CDEPPEVKNKPYSLFVPTSPSYELDEKALPQISENVIAWKRAFF 64
 QY 54 ESNVKNLISNLSDFIVKGMNKSAPFRBDVSFTFCANETMKSRTVM 124
 DB 54 ESNVKNLISNLSDFIVKGMNKSAPFRBDVSFTFCANETMKSRTVM 124
 QY 115 CANEMMGTALAVESDF 126
 DB 115 CANEMMGTALAVESDF 126

RESULT 4
 US 09 834 309 6 27
 Sequence 27, Application 05/007185, 670
 Patent No. 5816969
 GENERAL INFORMATION:

APPLICANT: Smith, Richard A.G.
 APPLICANT: Todd, Ian
 APPLICANT: Freeman, Mary A.
 APPLICANT: Messakowska, Danuta E.I.
 TITLE OF INVENTION: NO. 5816969 Compounds
 NUMBER OF SEQUENCES: 41
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: SmithKline Beecham Corporate Intellectual Property
 STREET: P.O. Box 1549
 CITY: King of Prussia
 STATE: Pennsylvania
 COUNTRY: USA
 ZIP: 19406
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 OPERATING SYSTEM: PC DOS/MS-DOS
 SOFTWARE: Patent In Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/456,461
 FILING DATE: 03-Jul-1995
 CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:
 NAME: Jervis, Herbert H.
 REGISTRATION NUMBER: 41,171
 REFERENCE/AGENT NUMBER: P0424
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (610) 270-5019
 TELEFAX: (610) 270-5090
 INFORMATION FOR SEQ ID NO: 27:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 197 amino acids
 TYPE: amino acid
 TOPOLOGY: linear
 MOLECULE TYPE: peptide
 FRAGMENT TYPE: N terminal
 US-08-456-461-27

Query Match 19.1% Score 144; DB 2; Length 197;
 Best Local Similarity 29.5% Pred. No. 43, 08;
 Matches 41; Conservative 22; Mismatches 54; Indels 22; Gaps 9;

QY 3 SCDEPPEVKNKPYSLFVPTSPSYELDEKALPQISENVIAWKRAFF 62
 DB 3 SCDEPPEVKNKPYSLFVPTSPSYELDEKALPQISENVIAWKRAFF 62
 QY 64 ESNVKNLISNLSDFIVKGMNKSAPFRBDVSFTFCANETMKSRTVM 114
 DB 64 ESNVKNLISNLSDFIVKGMNKSAPFRBDVSFTFCANETMKSRTVM 114
 QY 121 CD... PTHGLPPTITNGDFIST NRENTYGSVVTYRNGSGGKVVFLVGEPSLYC 175
 DB 121 CD... PTHGLPPTITNGDFIST NRENTYGSVVTYRNGSGGKVVFLVGEPSLYC 175
 QY 115 CANE... MW GPALPV 128
 DB 176 TSDQVGLWSGDA-PQ 192

RESULT 41
 US-08-769-967A 27
 Sequence 27, Application US/8876967A
 Patent No. 5859224
 GENERAL INFORMATION:

APPLICANT: Messakowska, Danuta E.I.
 APPLICANT: Smith, Richard A.G.
 APPLICANT: Todd, Ian
 APPLICANT: Freeman, Anne Mary
 TITLE OF INVENTION: Soluble CR1 Derivatives
 NUMBER OF SEQUENCES: 44
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: SmithKline Beecham Corporate Intellectual Property
 STREET: P.O. Box 1549
 CITY: King of Prussia
 STATE: Pennsylvania
 COUNTRY: USA
 ZIP: 19406

Genature version 5.1.1.4
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM protein protein search, using sw model
Run on: November 6, 2002 16:04:59 : Search time 14.5539 Seconds
(without alignments)
891.311 Million cell updates/sec

Title: US 09-844-409 6

Perfect score: 76.2

Sequence: 1 EISCDPEPEVKNAKPYSL.....ANEMWGTPALPVQESDPLE 145

Scoring table: BLOSUM62

Gapop 10.0, Gapext 0.5

Searched: 28148 seqs, 96089344 residues

Total number of hits satisfying chosen parameters: 281148

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Fast processing: Minimum Match 9%

Maximum Match 100%

Listing first 100 summaries

Database: PIR_71.*

1: pir1.*

2: pir2.*

3: pir3.*

4: pir4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	75.2	100.0	1025	1 A14526	complement C3d/eps
2	74.9	99.6	363	2 B45903	complement C3d/eps
3	74.9	99.6	676	2 A45909	complement C3b rec
4	466.5	62.0	1091	1 P10009	complement C3d/eps
5	242	40.9	597	1 S53711	C4b alpha chain p
6	225.5	30.0	597	1 N8H04	C4b-binding protei
7	222	29.5	469	1 N8H04	C4b-binding protei
8	221.5	29.5	2014	2 I36946	complement recepto
9	221.5	29.5	2489	2 I74012	complement C3b/C4b
10	215	28.6	579	2 A56740	sperm-eqg recognit
11	208.5	27.7	558	2 S57953	C4b protein alpha
12	200	26.6	610	1 I46001	C4b-binding protei
13	197.5	26.3	263	1 C36848	complement control
14	197.5	26.3	263	1 C36848	hypothetical prote
15	196.5	26.0	263	2 B2152	B181 protein var
16	194.5	25.7	263	1 WNVZSP	apolipoprotein h
17	189.5	25.1	449	2 G02913	sperm C4b - human
18	188.5	25.1	369	2 I57908	membrane cofactor
19	188.5	25.1	377	2 I54479	membrane cofactor
20	188.5	25.1	484	2 S01866	membrane cofactor
21	181.5	24.1	462	2 J05194	membrane cofactor
22	181.5	24.1	469	2 J05194	membrane cofactor
23	176	24.4	1244	1 N8M5H	complement factor
24	172.5	22.9	497	2 J02954	complement regulat
25	168	22.4	808	2 D45949	complement factor
26	167	22.2	449	2 A44516	complement recepto
27	164.5	21.9	482	2 A34948	complement C3b/C4b
28	157.5	20.9	669	2 S65551	factor II - bovine
29	156.5	20.8	252	2 A44877	C4b-binding protei

complement factor
X/Y protein - moos
complement factor
endothelial leukoc
complement factor
P-selectin precurs
secretory compleme
membrane bound com
decay-accelerating
decay-accelerating
P-selectin bean
apolipoprotein B p
complement control
E-selectin precurs
seizure-related pr
coagulation factor
coagulation factor
apolipoprotein B p
C4BP protein, beta
E-selectin precurs
scavenger receptor
complement C2 prec
complement factor
coagulation factor
decay-accelerating
P-selectin rat
apolipoprotein B p
complement factor
complement factor
P-selectin precurs
apolipoprotein B p
E-selectin - pig
apolipoprotein B p
hypothetical prote
hypothetical prote
hypothetical prote
hypothetical prote
ps/hr protein - va
complement factor
complement C6 prec
ASP protein - var
hypothetical prote
hypothetical prote
hypothetical prote
B2R protein - vari
B2R protein - vari
hypothetical prote
complement factor
F1V ap2, ps/hr pro
iodide peroxidase
iodide peroxidase
P-selectin precurs
complement factor
iodide peroxidase
P-selectin precurs
hypothetical prote
hypothetical prote
iodide peroxidase
C4BP beta chain
P-selectin precurs
P-selectin precurs
haptoglobin precurs

DB 256 P-ELPKC 261

RESULT 15
B72152

HHV protein, varicella minor virus (strain Garcia 1966)
C:Keywords: duplication; extracellular protein
F:119/Domain: signal sequence #status predicted -SLG-
F:20 26/Product: c4b binding protein homolog #status predicted -MAT-
F:21 81/Domain: complement factor H repeat homology -FH1-
F:86 144/Domain: complement factor H repeat homology -FH2-
F:148 201/Domain: complement factor H repeat homology -FH3-
F:206 261/Domain: complement factor H repeat homology -FH4-
C:Accession: B72152
C:Date: 24 Nov 1999 #sequence revision 24 Nov 1999 #text_change 20 Jun 2000
C:Author: Kikuchi, S.N., Tatematsu, A.V., Sutorov, V.V., Satomura, P.P., Messing, P.P.: Lysed
submitted to Genbank, March 1999
A:Description: Analysis of the complete coding sequence of DNA of a strain varicella minor
A:Reference number: A72150
A:Accession: B72152
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1 261 306
A:Cross-references: GB:V16780, N:159940055, P1DN:AA0461211, P1D:0580674
A:Experimental source: strain Garcia 1966
C:Status: confirmed

Query Match 25.78; Score 196.5; DB 1; Length 264;
Best Local Similarity 25.18; Prod. No. 256 11;
Matches 47; Conservative 19; Mismatches 60; Indels 11; Gaps 4;
QY 2 TSCDPPVKNARKKAYSLIVGIVLIVYVVSYSYVIGKALFETISENQVHATWKAPE 61
DB 146 VKQSPESISNRGNYDEYFVLSVVVYSNSGYSLLGNSGLTSG 36WSL PP 199
QY 6 TFSVVKTSLSLIVGIVGIMRSGKAPRHSVFTT-KANFTMKSGTIVWQANEMMG 121
DB 200 IYK -VKKVPTITLNGYLSGRKSYNNVDTCKYVYKLSGSSSSPSFSPENW 255
QY 122 PTALPVC 128
DB 256 P-ELPKC 261

RESULT 16
B72150

HHV protein, varicella minor virus (strain Garcia 1966)
C:Keywords: duplication; extracellular protein
F:119/Domain: signal sequence #status predicted -SLG-
F:20 26/Product: c4b binding protein homolog #status predicted -MAT-
F:21 81/Domain: complement factor H repeat homology -FH1-
F:86 144/Domain: complement factor H repeat homology -FH2-
F:148 201/Domain: complement factor H repeat homology -FH3-
F:206 261/Domain: complement factor H repeat homology -FH4-
C:Accession: B72150
C:Date: 24 Nov 1999 #sequence revision 24 Nov 1999 #text_change 20 Jun 2000
C:Author: Kikuchi, S.N., Tatematsu, A.V., Sutorov, V.V., Satomura, P.P., Messing, P.P.: Lysed
submitted to Genbank, March 1999
A:Description: Analysis of the complete coding sequence of DNA of a strain varicella minor
A:Reference number: A72150
A:Accession: B72152
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1 261 306
A:Cross-references: GB:V16780, N:159940055, P1DN:AA0461211, P1D:0580674
A:Experimental source: strain Garcia 1966
C:Status: confirmed

Query Match 25.78; Score 196.5; DB 2; Length 264;
Best Local Similarity 25.18; Prod. No. 256 11;
Matches 47; Conservative 19; Mismatches 60; Indels 11; Gaps 4;
QY 2 TSCDPPVKNARKKAYSLIVGIVLIVYVVSYSYVIGKALFETISENQVHATWKAPE 61
DB 146 VKQSPESISNRGNYDEYFVLSVVVYSNSGYSLLGNSGLTSG 36WSL PP 199
QY 6 TFSVVKTSLSLIVGIVGIMRSGKAPRHSVFTT-KANFTMKSGTIVWQANEMMG 121
DB 200 IYK -VKKVPTITLNGYLSGRKSYNNVDTCKYVYKLSGSSSSPSFSPENW 255
QY 122 PTALPVC 128
DB 256 P-ELPKC 261

RESULT 17
G02913

HHV protein, varicella minor virus (strain Garcia 1966)
C:Keywords: duplication; extracellular protein
F:119/Domain: signal sequence #status predicted -SLG-
F:20 26/Product: c4b binding protein homolog #status predicted -MAT-
F:21 81/Domain: complement factor H repeat homology -FH1-
F:86 144/Domain: complement factor H repeat homology -FH2-
F:148 201/Domain: complement factor H repeat homology -FH3-
F:206 261/Domain: complement factor H repeat homology -FH4-
C:Accession: G02913
C:Date: 21 Dec 1996 #sequence revision 06 Jun 1997 #text_change 20 Jun 2000
C:Author: Richard, T.
submitted to Genbank, March 1996
A:Reference number: H01942
A:Accession: G02913
A:Status: preliminary; translated from GH/EMM/00001
A:Molecule type: mRNA
A:Residues: 1 349 388
A:Cross-references: GB:D84105, N11041256200, P1DN:AA122411, P1D:01256201
C:Superfamily: herpesvirus complement control protein; complement factor H repeat hom
F:35 94/Domain: complement factor H repeat homology -FH01-
F:99 157/Domain: complement factor H repeat homology -FH02-
F:142-224/Domain: complement factor H repeat homology -FH03-
F:228-283/Domain: complement factor H repeat homology -FH04-
C:Status: confirmed

Query Match 25.18; Score 186.5; DB 2; Length 349;
Best Local Similarity 30.68; Prod. No. 196 10;
Matches 41; Conservative 24; Mismatches 55; Indels 15; Gaps 6;
QY 1 ELISLPPEVKNARKKAYSLIVGIVLIVYVVSYSYVIGKALFETISENQVHATWKAPE 64
DB 159 KVLTPPKIKENKRIITSEVVEVFLVAVVYSNPAPRPFESLISLIV 215
QY 55 TWKAPPICTESVVKITCSDFIVAGTGMNKKKALFEMASVDFKANKEMSKAVW 114
DB 216 -WSKAPPE - KVKRKEPVVVENKALISGPKPKPKYKALVMEPEKSFYLLSSSLIV 276
QY 115 -KARESWPTALIV 128
DB 271 DSNSLWGP PVKPS 283

RESULT 18
157998

membrane cofactor protein - human
C:Keywords: Homo sapiens (man)
C:Date: 22 Jul 1995 #sequence revision 22 Jul 1995 #text_change 10 May 2000
C:Accession: 157998
C:Author: F.J. Pauletti, P.; Akhondlou, C.; Bisi, R.; Rossi, R.
Mol. Reprod. Dev. 34: 107-118, 1993
A:Title: Characterization of a cDNA clone coding for human factor H membrane cofactor p
A:Reference number: 157998; MIMD:94119658

A:Accession: F57998
A:Status: preliminary; translated from GH/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-69-RES.
A:Cross-references: GB:51940; NID:q4262947; PIRN:AA84822.1; PDB:4z62938
C:Superfamily: herpesvirus complement control protein; complement factor H repeat homolo
F:35-94/Domain: complement factor H repeat homology <PH01>
F:99-157/Domain: complement factor H repeat homology <PH02>
F:162-224/Domain: complement factor H repeat homology <PH03>
F:228-284/Domain: complement factor H repeat homology <PH04>
Query Match 25.1% Score 188.5; DB 2; Length 469;
Best Local Similarity 30.6%; Pres. No. 1,86-10;
Matches 41; Conservative 23; Mismatches 55; Gaps 6;
QY 1 ELSTDPPEVNAKPKYYSSTIPVQ TVIPYTSPTS-----YPLIGKATPCTSENVVHA 54
DB 159 KVLCTPEPKIKNGKITSEVEVEVAVVTSVDPAPSPGPFSLIGSTIYC-GDNSV-- 215
QY 55 IWDKAPPTCESVNKLTSSTPVSQGMNKGSKAPFRKHGSVTFCKANFTMGSKIVWC 114
DB 216 WSPAAPPE---KVKKPEFVVENKGLSGPKKKFYKATVMEFCKGPFYLGSDPTIVC 270
QY 115 QANEMWHTALPVC 128
DB 271 DSNSTWDP-IVPKC 283
RESULT 19
154479
membrane cofactor protein precursor, splice form pm5.1 - human
N:Alternate names: lymphocyte surface glycoprotein CD46
C:Species: Homo sapiens (man)
C:Date: 07 Jun 1996 #sequence_revision 07 Jun 1996 #text_change 31-Mar-2000
A:Accession: F54479
A:Title: Alternatively spliced RNAs encode several isoforms of CD46 (MCP), a regulator o
A:Status: translated from GH/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-477-RES.
A:Cross-references: GH:F58050; NID:q180136; PIRN:AA62833.1; PDB:g180137
C:Genetics:
A:Gene: GDB:MCP
A:Cross-references: GH:120169; OMIM:120920
A:Map position: 1q42-1q42
C:Function:
A:Description: for the factor I mediated cleavage of the complement convertases
C:Superfamily: herpesvirus complement control protein; complement factor H repeat homolo
C:Keywords: alternative splicing; glycoprotein; transmembrane protein
F:35-94/Domain: complement factor H repeat homology <PH01>
F:99-157/Domain: complement factor H repeat homology <PH02>
F:162-224/Domain: complement factor H repeat homology <PH03>
F:228-284/Domain: complement factor H repeat homology <PH04>
F:329-351/Domain: transmembrane #status predicted <TM>
Query Match 25.1% Score 188.5; DB 2; Length 377;
Best Local Similarity 30.6%; Pres. No. 1,86-10;
Matches 41; Conservative 23; Mismatches 55; Gaps 6;
QY 1 ELSTDPPEVNAKPKYYSSTIPVQ TVIPYTSPTS-----YPLIGKATPCTSENVVHA 54
DB 159 KVLCTPEPKIKNGKITSEVEVEVAVVTSVDPAPSPGPFSLIGSTIYC-GDNSV-- 215
QY 55 IWDKAPPTCESVNKLTSSTPVSQGMNKGSKAPFRKHGSVTFCKANFTMGSKIVWC 114
DB 216 WSPAAPPE---KVKKPEFVVENKGLSGPKKKFYKATVMEFCKGPFYLGSDPTIVC 270
QY 115 QANEMWHTALPVC 128
DB 271 DSNSTWDP-IVPKC 283

RESULT 20
S01896
membrane cofactor protein precursor - human
N:Alternate names: lymphocyte surface glycoprotein CD46
C:Species: Homo sapiens (man)
C:Date: 30 Sep 1989 #sequence_revision 10-Sep-1989 #text_change 21-Jul-2000
A:Accession: S01896; A60765; 156188
R:Lablin, D.M.; Liszewski, M.K.; Post, T.W.; Allen, M.A.; Le Beau, M.M.; Redentzsch, M
J. Exp. Med. 168, 181-194, 1988
A:Title: Molecular cloning and chromosomal localization of human membrane cofactor pr
A:Reference number: S01896; M01D:88246C8C
A:Accession: S01896
A:Molecule type: mRNA
A:Residues: 1-384-CLB>
A:Cross-references: EMBL:Y00651; NID:q4504; PIRN:CA68675.1; PDB:34506
A:Note: part of this sequence, including the amino end of the mature protein, was con
R:Purell, D.F.J.; Deacon, N.J.; Andrew, S.M.; McKenzie, J.F.C.
Immunogenetics 31, 21-28, 1990
A:Title: Human non-lineage antigen, CD46 (LuLy m5): purification and partial sequenc
A:Reference number: A60765; M01D:50129152
A:Accession: A60765
A:Molecule type: protein
A:Residues: 'X'; 35-39; 'X'; 41-58 <PH0>
R:Cat, W.; Montecio, D.; Post, T.; Greenwood, A.C.; Atkinson, J.P.; Kimai, V
J. Immunol. 151, 4137-4146, 1993
A:Title: Characterization of the promoter region of the membrane cofactor protein (CD
A:Reference number: 156188; M01D:9401456
A:Accession: 156188
A:Status: preliminary; translated from 3H7MH/140BJ
A:Molecule type: DNA
A:Residues: 1-14 <RES>
A:Cross-references: GH:S65879; NID:q45043; PIRN:AA31068.1; PDB:31261068
C:Genetics:
A:Gene: GDB:MCP
A:Cross-references: GH:120169; OMIM:120920
A:Map position: 1q42-1q42
C:Function:
A:Description: membrane cofactor protein for the factor I-mediated cleavage of the co
C:Superfamily: herpesvirus complement control protein; complement factor H repeat hom
C:Keywords: alternative splicing; glycoprotein; transmembrane protein
F:1-34/Domain: signal sequence #status predicted <SIG>
F:35-384/product: membrane cofactor #status experimental <MA>
F:35-94/Domain: complement factor H repeat homology <PH01>
F:99-157/Domain: complement factor H repeat homology <PH02>
F:162-224/Domain: complement factor H repeat homology <PH03>
F:228-284/Domain: complement factor H repeat homology <PH04>
F:329-351/Domain: transmembrane #status predicted <TM>
Query Match 25.1% Score 188.5; DB 2; Length 484;
Best Local Similarity 30.6%; Pres. No. 1,86-10;
Matches 41; Conservative 23; Mismatches 55; Gaps 6;
QY 1 ELSTDPPEVNAKPKYYSSTIPVQ TVIPYTSPTS-----YPLIGKATPCTSENVVHA 54
DB 159 KVLCTPEPKIKNGKITSEVEVEVAVVTSVDPAPSPGPFSLIGSTIYC-GDNSV-- 215
QY 55 IWDKAPPTCESVNKLTSSTPVSQGMNKGSKAPFRKHGSVTFCKANFTMGSKIVWC 114
DB 216 WSPAAPPE---KVKKPEFVVENKGLSGPKKKFYKATVMEFCKGPFYLGSDPTIVC 270
QY 115 QANEMWHTALPVC 128
DB 271 DSNSTWDP-IVPKC 283
RESULT 21
305194
membrane cofactor protein precursor - atom monkey
N:Alternate names: MCP; measles virus receptor
C:Species: Cercopithecus aethiops (atom monkey) (qv) (qv)
C:Date: 20 Feb 1997 #sequence_revision 20 Feb 1997 #text_change 19 May 2000

Query Match 18.9%; Score 142; DB 1; Length 381;
Best Local Similarity 29.9%; Pred. No. 4.9e-06;
Matches 40; Conservative 24; Mismatches 50; Indels 20; Gaps 9;

Genware version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 6, 2002, 16:04:58 : Search time 8.5316 Seconds
(without alignments)
612.680 Million cell updates/sec

Title: US 09-834-309-6

Perfect score: 752

Sequence: 1 EISDPPEVKNARKYYSLS.....ANEMGPTALPVCEDFPLE 135

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 105224 seqs, 4870550 residues

Total number of hits satisfying chosen parameters: 105224

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post processing: Minimum Match 100

Maximum Match 100%

Listing first 100 summaries

Database : SwissProt_40:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query %	Match	Length	DB ID	Description
1	752	100.0	1025	1	C62_MOUSE	P19070 mus musculus
2	466.5	62.0	1033	1	C62_HUMAN	P30023 homo sapien
3	225.5	40.0	597	1	C4BP_HUMAN	P04003 homo sapien
4	222	29.5	469	1	C4BP_MOUSE	P08607 mus musculus
5	220.5	29.3	2039	1	C1L_HUMAN	P17927 homo sapien
6	208.5	27.7	558	1	C4BP_RAT	Q63514 rattus norv
7	200	26.6	610	1	C4BP_BOVIN	Q28065 bos taurus
8	193.5	25.7	263	1	VCP_VACCV	P10998 vaccinia vi
9	188.5	25.1	377	1	MP_HUMAN	P15529 homo sapien
10	176	23.4	1234	1	C7AH_MOUSE	P06909 mus musculus
11	157.5	20.9	585	1	C7AH_BOVIN	Q28085 bos taurus
12	156.5	20.8	252	1	C4AR_HUMAN	P29851 homo sapien
13	153.5	20.4	407	1	DAF2_MOUSE	Q61476 mus musculus
14	153.5	20.4	764	1	C7AB_HUMAN	P06751 homo sapien
15	149	19.8	1241	1	C7AB_HUMAN	P08603 homo sapien
16	146.5	19.7	551	1	LEM2_EBIBT	P27113 oryctolagus
17	148	19.7	390	1	DAF1_MOUSE	Q61475 mus musculus
18	144.5	19.2	151	1	C7AB_PIG	Q03710 sus scrofa
19	143.5	19.1	768	1	LEM3_MOUSE	Q01102 mus musculus
20	142.5	18.9	360	1	C7PI_HSVSA	Q01016 herpesvirus
21	142	18.9	381	1	DAF_HUMAN	P08174 homo sapien
22	142	18.9	485	1	LEM2_BOVIN	P98107 bos sapien
23	140	18.6	549	1	LEM2_RAT	P98105 rattus norv
24	139	18.5	610	1	LEM2_HUMAN	P16581 homo sapien
25	137	18.2	561	1	F13B_HUMAN	P05160 homo sapien
26	135	18.0	668	1	F13B_MOUSE	Q07968 mus musculus
27	134	17.8	297	1	APCH_RAT	P26644 rattus norv
28	133.5	17.8	258	1	C4BB_RAT	Q63515 rattus norv
29	131.5	17.5	612	1	LEM2_MOUSE	Q00690 mus musculus
30	130.5	17.4	752	1	C62_HUMAN	P06681 homo sapien
31	130.5	17.4	761	1	C7AB_MOUSE	P04186 mus musculus
32	130.5	17.4	1019	1	LPC_TACTR	P28175 tachypleus
33	130	17.3	340	1	DAF_PONPY	P49457 pongo pygma

34	129.5	17.2	768	1	LEM3_RAT	P94306 rattus norv
35	129	17.2	611	1	LEM2_CANIS	P43240 canis famill
36	128.5	17.1	769	1	LEM4_SHEEP	P94109 ovis aries
37	127.5	17.0	1019	1	LPC_CARRI	Q23422 carotinoscor
38	127	16.9	345	1	APCH_BOVIN	P17690 bos taurus
39	125.5	16.7	330	1	FHR1_HUMAN	Q03591 homo sapien
40	124.5	16.6	646	1	LEM3_BOVIN	P42201 bos taurus
41	124	16.5	345	1	AFCH_MOUSE	Q01335 mus musculus
42	123.5	16.4	484	1	LEM2_PIG	P98110 sus scrofa
43	123	16.4	345	1	APCH_CANIS	P43703 canis famill
44	121	16.1	507	1	DAF_CAVBO	Q60401 cavia porce
45	120	16.0	345	1	APCH_HUMAN	P02749 homo sapien
46	119	15.8	830	1	LEM3_HUMAN	P16109 homo sapien
47	118	15.7	958	1	HIG_DROME	Q09101 drosophila
48	117.5	15.6	202	1	APAR_MOUSE	Q03472 sus scrofa
49	117	15.6	760	1	C62_MOUSE	P21180 mus musculus
50	116	15.4	270	1	FHR2_HUMAN	P16480 homo sapien
51	113.5	15.1	843	1	C07_HUMAN	P10443 homo sapien
52	108.5	14.4	317	1	VB05_VACCV	Q11227 vaccinia vi
53	105.5	14.0	330	1	FHR3_HUMAN	Q02965 homo sapien
54	104.5	13.9	317	1	VB05_VACCV	P24084 vaccinia vi
55	104.5	13.9	317	1	VB05_VACCV	P24084 vaccinia vi
56	104.5	13.9	934	1	C6B_HUMAN	P13671 homo sapien
57	102.5	13.6	317	1	VB05_VACCV	P21115 vaccinia vi
58	100	13.3	464	1	SRPX_RAT	Q63769 rattus norv
59	99.5	13.2	250	1	CFBL_CUIR	P81475 gallus gall
60	99.5	13.2	331	1	FHR4_HUMAN	Q92496 homo sapien
61	97	12.9	464	1	SRPX_HUMAN	P78539 homo sapien
62	95.5	12.7	914	1	PERT_MOUSE	P45439 mus musculus
63	95.5	12.7	914	1	PERT_RAT	P44650 rattus norv
64	94.5	12.6	372	1	LEM1_PAT	P08046 rattus norv
65	93.5	12.4	933	1	PERT_HUMAN	P07202 homo sapien
66	88.5	11.8	372	1	LEM1_MACMU	Q95198 macaca mola
67	88.5	11.8	372	1	LEM1_PALFA	Q28768 papio hamad
68	87.5	11.6	926	1	PERT_PIC	P09933 sus scrofa
69	86	11.4	268	1	IL2A_CANIS	Q02802 canis famill
70	85	11.3	198	1	C4BB_BOVIN	Q28066 bos taurus
71	84.5	11.2	370	1	LEM1_BOVIN	P81311 bos taurus
72	84.5	11.2	372	1	LEM1_HUMAN	P14151 homo sapien
73	84.5	11.2	372	1	LEM1_PANR	Q5247 pan troglod
74	83.5	11.1	372	1	LEM1_PORLY	Q5245 pongo pygma
75	83	11.0	347	1	HFT_MOUSE	Q02558 mus saxicola
76	83	11.0	686	1	MAS2_HUMAN	G00147 homo sapien
77	80.5	10.7	268	1	IL2A_MOUSE	P31530 mus musculus
78	80.5	10.7	704	1	CHAR_MOUSE	P98364 mus musculus
79	79	10.5	275	1	IL2A_ELEA	P41640 felis silve
80	77.5	10.3	347	1	HPT_RAT	P06846 rattus norv
81	77.5	10.3	372	1	LEM1_MOUSE	P18347 mus musculus
82	76.5	10.2	699	1	CHAR_HUMAN	P48740 h. complimen
83	76	10.1	1248	1	DIA1_HUMAN	G00610 homo sapien
84	75.5	10.0	1627	1	PAPA_HUMAN	Q13219 homo sapien
85	74.5	9.9	272	1	IL2A_HUMAN	P01589 homo sapien
86	74	9.9	705	1	C1P_HUMAN	P00736 homo sapien
87	72.5	9.6	1257	1	PCCN_RAT	P55067 rattus norv
88	72	9.6	347	1	HFT_ATRIS	P00417 athera aequi
89	72	9.6	745	1	ZF2_HUMAN	Q05996 homo sapien
90	72	9.6	960	1	GHR1_MOUSE	G99018 mus musculus
91	72	9.6	961	1	GHR1_HUMAN	Q90854 homo sapien
92	72	9.6	991	1	GHR1_RAT	G90004 rattus norv
93	72	9.6	1255	1	DAL_MOUSE	G08008 mus musculus
94	71	9.4	291	1	DM43_DIANA	P82957 diadophis m
95	71	9.4	3381	1	PCCV_BOVIN	P81282 bos taurus
96	70.5	9.4	651	1	Y942_METJA	Q58452 metatherocon
97	70.5	9.4	837	1	MUC1_RAT	P98089 rattus norv
98	70	9.3	666	1	ZP2_LABIT	P48829 c. talassus
99	70	9.3	745	1	ZP2_NACCA	Q77243 nassa radi
100	70	9.3	839	1	PHLD_BOVIN	P80109 bos taurus

ALIGNMENTS

RESULT 1

FT D1S0LFD 458 400 HY SIMILARITY.
 FT D1S0LFD 486 416 HY SIMILARITY.
 FT D1S0LFD 421 470 HY SIMILARITY.
 FT D1S0LFD 450 487 HY SIMILARITY.
 FT D1S0LFD 494 546 HY SIMILARITY.
 FT D1S0LFD 524 549 HY SIMILARITY.
 FT D1S0LFD 554 595 HY SIMILARITY.
 FT D1S0LFD 581 611 HY SIMILARITY.
 FT D1S0LFD 616 665 HY SIMILARITY.
 FT D1S0LFD 645 682 HY SIMILARITY.
 FT D1S0LFD 684 740 HY SIMILARITY.
 FT D1S0LFD 716 743 HY SIMILARITY.
 FT D1S0LFD 747 790 HY SIMILARITY.
 FT D1S0LFD 776 803 HY SIMILARITY.
 FT D1S0LFD 804 850 HY SIMILARITY.
 FT D1S0LFD 846 866 HY SIMILARITY.
 FT D1S0LFD 871 920 HY SIMILARITY.
 FT D1S0LFD 903 937 HY SIMILARITY.
 FT D1S0LFD 943 986 HY SIMILARITY.
 FT D1S0LFD 974 999 HY SIMILARITY.
 FT D1S0LFD 1004 1045 HY SIMILARITY.
 FT D1S0LFD 1041 1061 HY SIMILARITY.
 FT D1S0LFD 1066 1115 HY SIMILARITY.
 FT D1S0LFD 1095 1132 HY SIMILARITY.
 FT D1S0LFD 1138 1180 HY SIMILARITY.
 FT D1S0LFD 1166 1193 HY SIMILARITY.
 FT D1S0LFD 1197 1240 HY SIMILARITY.
 FT D1S0LFD 1226 1253 HY SIMILARITY.
 FT D1S0LFD 1258 1300 HY SIMILARITY.
 FT D1S0LFD 1286 1316 HY SIMILARITY.
 FT D1S0LFD 1321 1370 HY SIMILARITY.
 FT D1S0LFD 1350 1387 HY SIMILARITY.
 FT D1S0LFD 1396 1439 HY SIMILARITY.
 FT D1S0LFD 1426 1452 HY SIMILARITY.
 FT D1S0LFD 1457 1498 HY SIMILARITY.
 FT D1S0LFD 1484 1514 HY SIMILARITY.
 FT D1S0LFD 1519 1568 HY SIMILARITY.
 FT D1S0LFD 1548 1585 HY SIMILARITY.
 FT D1S0LFD 1591 1633 HY SIMILARITY.
 FT D1S0LFD 1619 1646 HY SIMILARITY.
 FT D1S0LFD 1650 1693 HY SIMILARITY.
 FT D1S0LFD 1679 1706 HY SIMILARITY.
 FT D1S0LFD 1711 1753 HY SIMILARITY.
 FT D1S0LFD 1739 1769 HY SIMILARITY.
 FT D1S0LFD 1774 1823 HY SIMILARITY.
 FT D1S0LFD 1803 1840 HY SIMILARITY.
 FT D1S0LFD 1848 1891 HY SIMILARITY.
 FT D1S0LFD 1877 1904 HY SIMILARITY.
 FT D1S0LFD 1909 1952 HY SIMILARITY.
 FT D1S0LFD 1938 1965 HY SIMILARITY.
 FT C4BP-RAT 56 56 N-LINKED (GLCNAC...) (POTENTIAL).
 FT C4BP-RAT 252 252 N-LINKED (GLCNAC...) (POTENTIAL).
 FT C4BP-RAT 410 410 N-LINKED (GLCNAC...) (POTENTIAL).
 FT C4BP-RAT 447 447 N-LINKED (GLCNAC...) (POTENTIAL).
 FT C4BP-RAT 509 509 N-LINKED (GLCNAC...) (POTENTIAL).
 FT C4BP-RAT 578 578 N-LINKED (GLCNAC...) (POTENTIAL).
 FT C4BP-RAT 702 702 N-LINKED (GLCNAC...) (POTENTIAL).
 FT C4BP-RAT 860 860 N-LINKED (GLCNAC...) (POTENTIAL).
 FT C4BP-RAT 897 897 N-LINKED (GLCNAC...) (POTENTIAL).
 FT C4BP-RAT 959 959 N-LINKED (GLCNAC...) (POTENTIAL).
 FT C4BP-RAT 1028 1028 N-LINKED (GLCNAC...) (POTENTIAL).
 FT C4BP-RAT 1152 1152 N-LINKED (GLCNAC...) (POTENTIAL).
 FT C4BP-RAT 1310 1310 N-LINKED (GLCNAC...) (POTENTIAL).
 FT C4BP-RAT 1481 1481 N-LINKED (GLCNAC...) (POTENTIAL).
 FT C4BP-RAT 1504 1504 N-LINKED (GLCNAC...) (POTENTIAL).
 FT C4BP-RAT 1544 1544 N-LINKED (GLCNAC...) (POTENTIAL).
 FT C4BP-RAT 1540 1540 N-LINKED (GLCNAC...) (POTENTIAL).
 FT C4BP-RAT 1605 1605 N-LINKED (GLCNAC...) (POTENTIAL).
 FT C4BP-RAT 1763 1763 N-LINKED (GLCNAC...) (POTENTIAL).
 FT C4BP-RAT 1908 1908 N-LINKED (GLCNAC...) (POTENTIAL).
 FT C4BP-RAT 2039 2039 N-LINKED (GLCNAC...) (POTENTIAL).
 SQ SEQUENCE AA; 224599 MW; B2FD29B6AD3C5EB7 CRC64;

Query Match

29.3% score 220.5; DB 1; Length 2039;

Best local Similarity 36.0%; Pred. No. 4,1e 14;
 Matches 49; Conservative 18; Mismatches 54; Indels 15; Gaps 5;
 QY 2 ISCDPP-EVKNARKPYYS---LPVPGTVLYT-SPS-----YRLGKKAIPGISENV 5;
 DB 1517 ISCPDPPTISNG--DPYSNNRTSPHNTVVIVYV-AGDGD-SDQLEFELVGGESVYVSKIDQ 1574
 QY 53 HATWKAIP-ICESVKNK-ISCSDPILVPGGEMKNK-SSAIPK-VDVDF-CAANFIMKSKIV 112
 DB 1575 VGVSSPPPKICSTNK---CTAPEVENAIRVPGKVSFSLTEIIRKCPGFGVWGSHTV 1641
 QY 113 WCOANEMMCPTALPVC 128
 DB 1642 OCQTNGKWP-KLPKRC 1646
 RESULT 6
 C4BP-RAT ID C4BP-RAT STANDARD; PRT; 533 AA.
 AC Q65514;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE C4B-binding protein alpha chain precursor (c4bp).
 GN C4BPA.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Granata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria, Rodentia; Sciurognathia; Muridae; Rattus.
 OX NCBI TaxID=10116;
 RN [1]
 SEQUENCE FROM N.A.
 RC STRAIN-SPRAGUE-DAWLEY; TISSUE-Liver;
 RX MEELINE-971f6082; Pubmed-9013975;
 RA Hillarp A., Wiklund H., Thern A., Dahlback B.
 FT "Molecular cloning of rat C4b binding protein alpha- and beta-chains:
 FT structural and functional relationships among human, bovine, rabbit
 FT mouse, and rat proteins";
 RC J. Immunol. 158:1315-1323(1997).
 CC -!- FUNCTION: C4BP CONTROLS THE CLASSICAL PATHWAY OF COMPLEMENT
 CC ACTIVATION. IT BINDS AS A COFACTOR TO C3b/C4b INACTIVATOR
 CC (C3bINA), WHICH THEN HYDROLYSES THE COMPLEMENT FRAGMENT C4b. IT
 CC ALSO ACCELERATES THE DEGRADATION OF THE C4b2a COMPLEX (C4
 CC CONVERTASE) BY DISSOCIATING THE COMPLEMENT FRAGMENT C2a. ALPHA
 CC CHAIN BINDS C4b. IT INTERACTS ALSO WITH ANTICOAGULANT PROTEIN S
 CC AND WITH SERUM AMYLOID P COMPONENT
 CC -!- SUBUNIT: DISULFIDE-LINKED COMPLEX OF C4BP ALPHA AND BETA CHAINS.
 CC -!- SIMILARITY: TO C4BP BETA CHAIN AND TO PIG APOLIPOPROTEIN R.
 CC -!- SIMILARITY: CONTAINS 8 SUSHI (SCX) DOMAINS.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by author commercial
 CC entities requires a license agreement. (see http://www.isb-sib.ch/about/)
 CC or send an email to license@isb-sib.ch).
 CC EMBL: Z50051; CAA90391.1;
 DR HSSP: P10998; LVVD.
 DR InterPro: IPR000436; Sushi_SCR_CCP.
 DR Pfam: PF00084; sushi; 8.
 DR SMART: SM00042; CCP; 8.
 KW Complement pathway; Plasma; Glycoprotein; Repeat; Sushi; Signal.
 FT SIGNAL 1 13 BY SIMILARITY.
 FT CHAIN 14 558 C4B-BINDING PROTEIN ALPHA CHAIN.
 FT DOMAIN 76 73 SUSHI 1.
 FT DOMAIN 14 135 SUSHI 2.
 FT DOMAIN 176 200 SUSHI 3.
 FT DOMAIN 203 259 SUSHI 4.
 FT DOMAIN 262 325 SUSHI 5.
 FT DOMAIN 328 387 SUSHI 6.
 FT DOMAIN 389 444 SUSHI 7.
 FT DOMAIN 446 502 SUSHI 8

A* p1099H.
 DT 01 JUL 1989 (Rel. 11, Created)
 DT 01 JUL 1989 (Rel. 11, Last sequence update)
 DT 16 OCT 2001 (Rel. 40, Last annotation update)
 DE Complement control protein precursor (MCP) (Secretory protein 35)
 DE (Protein C3) (28 kDa protein).
 GN C3L.
 OS Vaccinia virus (strain WR), and
 OS Vaccinia virus (strain Copenhagen).
 OC Viruses; dsDNA viruses, no RNA stage; Poxviridae; Chordopoxvirinae;
 OC orthopoxvirus.
 OX NEB_TaxID=10254, 10249;
 RN [1]
 RP SEQUENCE FROM N.A., AND SEQUENCE OF 20-37.
 RC STRAIN-WR.
 RX MEDLINE-BH318974; PubMed-4412473;
 RA Kotwal G.J., Moss B.;
 RT "Vaccinia virus encodes a secretory polypeptide structurally related
 RL to complement control proteins.";
 RL Nature 345:176-178(1988)
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN-WR.
 RX MEDLINE-B9074756; PubMed-2849248;
 RA Kotwal G.J., Moss B.;
 RT "Analysis of a large cluster of nonessential genes deleted from a
 RL vaccinia virus terminal transposition mutant.";
 RL Virology 167:524-547(1988).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN-COPENHAGEN;
 RX MEDLINE-Q1021027; PubMed-2219722;
 RA Goebel S.J., Johnson G.P., Perkins M.F., Davis S.W., Winslow J.P.,
 RA Paoletti E.;
 RT "The complete DNA sequence of vaccinia virus.";
 RL Virology 179:247-266(1990).
 RN [4]
 RP COMPLETE GENOME.
 RC STRAIN-COPENHAGEN.
 RA Goebel S.J., Johnson G.P., Perkins M.E., Davis S.W., Winslow J.P.,
 RA Paoletti E.;
 RT "Appendix to 'The complete DNA sequence of vaccinia virus'.";
 RL Virology 179:517-563(1990).
 RN [5]
 RP FUNCTION.
 RX MEDLINE-Q2115714; PubMed-1731333;
 RA Isaacs S.N., Kotwal G.J., Moss B.;
 RT "Vaccinia virus complement control protein prevents
 RL antibody-dependent complement-enhanced neutralization of infectivity
 RL and contributes to virulence.";
 RL Proc. Natl. Acad. Sci. U.S.A. 89:628-632(1992).
 RN [6]
 RP STRUCTURE BY NMR OF 146-264.
 RX MEDLINE-Q7446168; PubMed-9299452;
 RA Miles A.P., Shaw G., Bright J., Petrczel A., Campbell I.D.,
 RA Barlow P.N.;
 RT "NMR studies of a viral protein that mimics the regulators of
 RL complement activation.";
 RL J. Mol. Biol. 272:253-265(1997).
 RN [7]
 RP FUNCTION: SERVES TO PROTECT THE VIRUS AGAINST COMPLEMENT ATTACK BY
 RN INHIBITING BOTH CLASSICAL AND ALTERNATIVE PATHWAYS OF COMPLEMENT
 RN ACTIVATION, BINDS C3B AND C4B.
 RN [8]
 RP SIMILARITY: BELONGS TO THE SUPERFAMILY OF THE REGULATORS OF
 RN COMPLEMENT ACTIVATION (SCRA).
 RN [9]
 RP SIMILARITY: CONTAINS 4 SUSHI (SCR) DOMAINS.
 RN [10]
 RP THIS SWISS PROTEIN ENTRY IS COPYRIGHT. IT IS PRODUCED THROUGH A COLLABORATION
 RN between the Swiss Institute of Bioinformatics and the EMBL outstation -
 RN the European Bioinformatics Institute. There are no restrictions on its
 RN use by non profit institutions as long as its content is in no way
 RN modified and this statement is not removed. Usage by and for commercial
 RN entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 RN or send an email to license@isb-sib.ch).

CC EMBL: X13156; CAA31564.1;
 DR EMBL: M22812; AAA69605.1;
 DR EMBL: M35027; AAA47997.1;
 DR EMBL: A31003; WWVZSP.
 DE PDB: 1VVC; 02-DEC-97.
 DR PDB: 1VVD; 02-DEC-97.
 DR PDB: 1VVE; 02-DEC-97.
 DR InterPro: IPK00436; Sushi_SCR_CCP
 DR Pfam: PF00384; sushi; 4.
 DR SMART: SMO10.2; CCP; 4.
 KW Signal; Repeat; Sushi; 3D-structure.
 FT SIGNAL: 1 19
 FT CHAIN: 20 263 COMPLEMENT CONTROL PROTEIN.
 FT DOMAIN: 20 82 SUSHI 1.
 FT DOMAIN: 85 144 SUSHI 2.
 FT DOMAIN: 147 202 SUSHI 3.
 FT DOMAIN: 205 262 SUSHI 4.
 FT DISULFID: 21 70 HY SIMILARITY.
 FT DISULFID: 54 81 HY SIMILARITY.
 FT DISULFID: 86 126 BY SIMILARITY.
 FT DISULFID: 112 143 BY SIMILARITY.
 FT DISULFID: 148 190 BY SIMILARITY.
 FT DISULFID: 176 201 HY SIMILARITY.
 FT DISULFID: 206 248 HY SIMILARITY.
 FT DISULFID: 234 261 HY SIMILARITY.
 SQ SEQUENCE 263 AA; 28629 MW; E4325C9A6EE8997 CRC64;
 Query Match 25.78; Score 19.5; PB.1; Length 263;
 Post local Similarity 29.18; Pred. No. 2.2; 12;
 Matches 37; Conservative 19; Mismatches 60; Indels 11; Gaps 4.
 QY 2 ISCTPPPPVFNARKPYSLPIVPTVTPYVWISSEYPLIRKKAIFQISENVAHATWRFAPP 61
 Db 146 VKQSPSPISNGKNGSYEDFYDASVVTYSNGYSLGNSGVLSG ---GEWSD-PP 199
 QY 62 ICKSVNKVISCSDPIVPGGPNKSKAPF-RHCI SVTFIRKANFTMKSKTVWQANHWG 121
 Db 200 TQQ-----VKPHPTISNGYSLSSGPKRSYSYN NVDFQKYGKLSGSSSTSPQNTWR 255
 QY 122 PTALFVC 128
 Db 256 P-ELPKC 261
 RESULT 9
 MCP_HUMAN
 ID MCP_HUMAN STANDARD; PRT; 277 AA.
 AC P15529;
 DT 01-APR-1990 (Rel. 14, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Membrane cofactor protein precursor (CD46 antigen) (Tropoelastat
 DE leucocyte common antigen) (ILX).
 GN MCP.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A., AND SEQUENCE OF 1-58.
 RX MEDLINE-B628080; PubMed-3260937;
 RA Lublin D.P., Liszewski M.K., Post T.V., Arco M.A., Le Beau M.M.,
 RA Reibel M.B., Lemons R.S., Seya T., Atkinson J.P.;
 RT "Molecular cloning and chromosomal localization of human membrane
 RL cofactor protein (MCP). Evidence for inclusion in the multigene
 RL J. Exp. Med. 168:181-194(1988).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC TISSUE-Testis;
 RX MEDLINE-Q3114658; PubMed-8418811;
 RA Cervoni F., Benichou P., Akhoundi S., Hsi B.L., Rossi R.;

FT DOMAIN 22 77
 FT DOMAIN 80 135
 FT DOMAIN 108 192
 FT DISULFID 23 63
 FT DISULFID 49 76
 FT DISULFID 81 121
 FT DISULFID 107 134
 FT DISULFID 139 179
 FT DISULFID 165 191
 FT DISULFID 202 202
 FT DISULFID 216 216
 FT CARBOHYD 64 64
 FT CARBOHYD 71 71
 FT CARBOHYD 98 98
 FT CARBOHYD 117 117
 FT CARBOHYD 154 154
 FT VARIANT 198 198
 FT SEQUENCE 25 AA: 28 457 MW: 6166664067052E7 CRC64;
 Query Match 20.8%; Score 156.5; DB 1; Length 252;
 Best Local Similarity 28.9%; Pred. No. 1.1e-08;
 Matches 47; Conservative 19; Mismatches 57; Indels 15; Gaps 6;
 QY 4 CPPPEVKNARKPYSLPVTGLRYTGSPSYRLGKAIKFCISENQVHATWDKAPPI 62
 DB 23 CPELPVNS---IFVAKEVEQILGTVGCIKGHYLVKKTLFCNASKE---WDNTTTE 75
 QY 63 QSVVKNKTSQSDPIVGGFMNKKAPRHGDSVTFCTKANFTMKGSKVWCQANEMWGP 122
 DB 76 C---RLGHCPOPVLVNGEFS--SSGPNVNSDKITFMNHYILKGSNSQJLEDHTWAP 129
 QY 124 TALPQVES 130
 DB 130 -PPPICKS 136
 RESULT 14
 LAF2 MOUSE STANDARD; PRT; 407 AA.
 AC Q51476;
 DT 01-NOV-1997 (rel. 35, Created)
 DT 01-NOV-1997 (rel. 35, Last sequence update)
 DT 01-NOV-1997 (rel. 35, Last annotation update)
 DE Complement decay-accelerating factor, Transmembrane precursor
 DE (LAF TM).
 GN DAF2.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN 111
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J; TISSUE=Testis;
 RX MEDLINE=93404982; PubMed=7345711;
 RA Spicer, A.P., Seldin M.F., Gender S.J.;
 RT *Molecular cloning and chromosomal localization of the mouse decay-
 RT accelerating factor genes, duplicated genes encode
 RI glycosylphosphatidylinositol anchored and transmembrane forms.*;
 RL J. Immunol. 155: 3079-3091 (1995).
 CC 1 FUNCTION: PROTECTION OF CELLS FROM COMPLEMENT-MEDIATED DAMAGE (BY
 CC SIMILARITY)
 CC 1 SUBCELLULAR LOCATION Type I membrane protein (Potential).
 CC 1 TISSUE SPECIFICITY TESTES, SPLEEN AND LYMPH NODE.
 CC 1 DOMAIN: THE FIRST SUSHI DOMAIN (SCR1) IS NOT NECESSARY FOR
 CC FUNCTION. SCR2 AND SCR4 PROVIDE THE PROPER CONFORMATION FOR THE
 CC ACTIVE SITE ON SCR3 (BY SIMILARITY).
 CC 1 SIMILARITY: CONTAINS 4 SUSHI (SCR) DOMAINS.
 CC 1 SIMILARITY: BELONGS TO THE RECEPTORS OF COMPLEMENT ACTIVATION
 CC (SCA) FAMILY.

CC This SWISS-PROT entry is copyright, it is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.sdb.ssi.ch>)
 CC or send an email to license@sdb.ssi.ch.
 DR EMBL: L41465; AAB00092.1;
 DR ISSP: P08603; IHCC.
 DR MGI: MGI:104849; Dat2.
 DR In-epro: IPR000436; Sush1_SCR_CCP.
 DR Pfam: PF00084; sush1; 4.
 DR SMART: SM00332; CCP; 4.
 KW Complement pathway; Glycoprotein; Repeat; Signal; Sushi;
 KW Transmembrane
 FT SIGNAL 1 39
 FT CHAIN 40 407
 FT DOMAIN 40 368
 FT TRANSMEM 369 389
 FT DOMAIN 390 407
 FT DOMAIN 40 100
 FT DOMAIN 102 164
 FT DOMAIN 167 226
 FT DOMAIN 229 290
 FT DOMAIN 291 363
 FT DISULFID 70 99
 FT DISULFID 133 150
 FT DISULFID 134 163
 FT DISULFID 158 209
 FT DISULFID 195 225
 FT DISULFID 230 272
 FT DISULFID 258 289
 FT CARBOHYD 132 192
 FT CARBOHYD 267 267
 SQ SEQUENCE 407 AA: 44469 MW: 600826372C6A40 CRC64;
 Query Match 20.4%; Score 153.5; DB 1; Length 407;
 Best Local Similarity 27.5%; Pred. No. 3.7e-08;
 Matches 39; Conservative 31; Mismatches 53; Indels 19; Gaps 8;
 QY 3 SCPPPEVKNARKPYSLP---IVPCTIVRYTGSYSYRLGKAIKFCISENQVHATWDK 59
 DB 157 SCPPNPKDLNG---HINPTGLGFGSEINSONGYRLGKITSILCTITGNV--IWIDE 221
 QY 50 PPICEVKNKTSQSDPIVGGFMNKKAPRHGDSVTFCTKANFTMKGSKVWCQANE 118
 DB 222 FVCTHE----FCDPPPKINDGIMKESDSYKSGQVTFISCDKGFIFGNSIFVCSKS 277
 QY 119 ---MWPTALPVC--ESDFPLE 135
 DB 278 DVGQW3-SPPQCLEESKVPK 298
 RESULT 14
 CFAE_HUMAN STANDARD; PRT; 764 AA
 AC P00751; O130(6); Q29944; Q9BTF5; Q9EX92;
 DT 2-JUL-1995 (rel. 01, Created)
 DT 0-OCT-1994 (rel. 30, Last sequence update)
 DT 16-OCT-2001 (rel. 40, Last annotation update)
 DE Complement factor B precursor (EC 3.4.21.4); (C3/C5 convertase)
 DE (properdin factor B) (Glycine-rich beta glycoprotein) (G3G) (FBF2)
 GN BF.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
 OX NCBI_TaxID=9606;
 RN 111
 RP SEQUENCE FROM N.A. (ALLELES S; FA AND FB).
 RX MEDLINE=91365702; PubMed=2249879;
 RA Davrinche C., Abbal M., Clere A.;

SEQUENCE FROM N.A.
 TISSUE=Heart;
 MEDLINE=92189729; PubMed=14721692;
 RA Lariquin J.D., Tsang J.C., Rumberger J.M., Harris D.K.;
 "Characterization of cDNA and genomic sequences encoding rabbit
 ELAM 1: conservation of structure and functional interactions with
 leukocytes.";
 RL DNA Cell Biol. 11:149-162(1992).
 CC ! FUNCTION: EXPRESSED ON CYTOKINE INDUCED ENDOTHELIAL CELLS AND
 CC MEDIATES THEIR BINDING TO LEUKOCYTES. THE LIGAND RECOGNIZED BY
 CC ELAM 1 IS STALK-LINKS X (ALPHA(1-3)FUCOSYLATED DERIVATIVES OF
 CC POLYGLUTAMINE THAT ARE FOUND AT THE NONREDUCING TERMINI OF
 CC GLYCOLIPIDS)
 CC ! SUBCELLULAR LOCATION: Type 1 membrane protein.
 CC ! INDUCTION: BY CYTOKINES.
 CC ! SIMILARITY: TO OTHER SELECTINS/LECAMS
 CC ! SIMILARITY: CONTAINS 1 C-TYPE LECTIN FAMILY DOMAIN.
 CC ! SIMILARITY: CONTAINS 1 EGF-LIKE DOMAIN
 CC ! SIMILARITY: CONTAINS 5 SUSHI (SCR) DOMAINS.
 CC
 CC THIS SWISS PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC
 DR EMBL: M91004; AAA31243.1;
 DR EMBL: M91005; AAA31244.1;
 DR ISSP: P16581; 187A.
 DR InterPro: IPR000561; EGF-like.
 DR InterPro: IPR002396; Selectin.
 DR InterPro: IPR000436; Sushi_SCR_NCP.
 DR InterPro: IPR001404; Lectin_C.
 DR Pfam: PF00008; EGF_1.
 DR Pfam: PF00059; Lectin_C_1.
 DR Pfam: PF00084; Sushi_5.
 DR PRINTS: PR00343; SELECTIN.
 DR SMART: SM00042; CYP_5.
 DR SMART: SM00044; CLECT_1.
 DR SMART: SM00181; EGF_1.
 DR PROSITE: PS00022; EGF_1; 1.
 DR PROSITE: PS01186; EGF_2; 1.
 DR PROSITE: PS00615; C_TYPE_LECTIN_1; 1.
 DR PROSITE: PS50043; C_TYPE_LECTIN_2; 1.
 KW Cell adhesion; Transmembrane; Glycoprotein; EGF-like domain; Lectin;
 FT SIGNAL 1 24
 FT CHAIN 24 551 E-SELECTIN.
 FT DOMAIN 24 445 EXTRACELLULAR (POTENTIAL).
 FT TRANSIMP 496 517 POTENTIAL.
 FT DOMAIN 518 551 CYTOPLASMIC (POTENTIAL).
 FT DOMAIN 40 140 C-TYPE LECTIN (SHORT FORM).
 FT DOMAIN 141 177 EGF-LIKE.
 FT DOMAIN 181 240 SUSHI_1.
 FT DOMAIN 243 302 SUSHI_2.
 FT DOMAIN 305 365 SUSHI_3.
 FT DOMAIN 368 428 SUSHI_4.
 FT DOMAIN 431 487 SUSHI_5.
 FT DISULFID 42 140 HY SIMILARITY.
 FT DISULFID 113 132 HY SIMILARITY.
 FT DISULFID 145 156 HY SIMILARITY.
 FT DISULFID 150 165 HY SIMILARITY.
 FT DISULFID 167 176 HY SIMILARITY.
 FT DISULFID 182 226 HY SIMILARITY.
 FT DISULFID 212 239 HY SIMILARITY.
 FT DISULFID 244 288 HY SIMILARITY.
 FT DISULFID 274 301 HY SIMILARITY.
 FT DISULFID 306 351 HY SIMILARITY.
 FT DISULFID 337 364 HY SIMILARITY.
 FT DISULFID 369 414 HY SIMILARITY.
 FT DISULFID 400 427 HY SIMILARITY.

FT DISULFID 432 473 BY SIMILARITY.
 FT DISULFID 459 486 BY SIMILARITY.
 FT CARBOHYD 32 32 N-LINKED (GLCNAC...) (POTENTIAL).
 FT CARBOHYD 45 45 N-LINKED (GLCNAC...) (POTENTIAL).
 FT CARBOHYD 162 162 N-LINKED (GLCNAC...) (POTENTIAL).
 FT CARBOHYD 194 194 N-LINKED (GLCNAC...) (POTENTIAL).
 FT CARBOHYD 201 201 N-LINKED (GLCNAC...) (POTENTIAL).
 FT CARBOHYD 205 205 N-LINKED (GLCNAC...) (POTENTIAL).
 FT CARBOHYD 267 267 N-LINKED (GLCNAC...) (POTENTIAL).
 FT CARBOHYD 314 314 N-LINKED (GLCNAC...) (POTENTIAL).
 FT CARBOHYD 321 321 N-LINKED (GLCNAC...) (POTENTIAL).
 FT CARBOHYD 334 334 N-LINKED (GLCNAC...) (POTENTIAL).
 FT CARBOHYD 442 442 N-LINKED (GLCNAC...) (POTENTIAL).
 FT CARBOHYD 466 466 N-LINKED (GLCNAC...) (POTENTIAL).
 FT CONFLICT 308 408 T > A (IN REF. 1; AAA31244).
 FT CONFLICT 328 428 T > A (IN REF. 1; AAA31244).
 FT CONFLICT 491 491 A > V (IN REF. 1; AAA31244).
 SQ SEQUENCE 551 AA; 60346 MW; 23B78A33B24240E GR064;
 Query Match 19.7%; Score 143.5; Db 1; Length 551;
 Best Local Similarity 25.9%; Pred No 1.6e-17;
 Matches 36, Conservative 20, Mismatches 48, Indels 35, Gaps 6;
 QY 1 ETSQDPEVKNRKPYSLSPIVGTVLRYTSPSYELIGEKALPQISENQVHATWIKAP 60
 DB 254 DVKCS-----SGSSAPW-----NTGTLDESEGFILGAKSLQTS-----GSMUNEK 298
 QY 61 PICES-----NKTISQSDPIVPGCFMNRKSKAFERHCHSVIFCKANFTMKSK 111
 DB 259 FTCKAVSLTIHMPONGSVGCSN-----SSGKTFPRSSCNFCEENLELRGPAQ 448
 QY 112 VMCOANEMWGPALPVCS 130
 DB 349 VECTAGLWFOQA-PVCHFA 366
 RESULT 17
 DAF1_MOUSE
 AC Q61475, Q61357, p97732, PRP; 340 AA.
 DT 01-NOV-1997 (Ref. 35, Created)
 DT 01-NOV-1997 (Ref. 35, Last sequence update)
 DT 15-JUL-1994 (Ref. 38, Last annotation update)
 DE Complement decay-accelerating factor, 3PI anchored precursor
 GN (DAF-GPI).
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sauriunguathia; Muridae; Mus.
 OC NCBI_TaxID=10990;
 RN 1;
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J; TISSUE=Testis;
 RX MEDLINE=95104982; PubMed=7545711;
 RA Spicer A.P., Seldin M.F., Gendler S.J.;
 RT "Molecular cloning and chromosomal localization of the mouse decay
 RT accelerating factor genes. duplicated genes encode
 RT glycosylphosphatidylinositol-anchored and transmembrane forms.";
 RL J. Immunol. 155:3079-3091(1995).
 RN [2]
 RP SEQUENCE OF 7-390 FROM N.A.
 RC STRAIN=BALB/c; TISSUE=Spleen;
 RX MEDLINE=9646213; PubMed=8671624;
 RA Fukunaka Y., Yasui A., Okada H.;
 RT "Molecular cloning of murine decay accelerating factor by
 RT immunoscreening.";
 RL Int. Immunol. 8:379-385(1996).
 CC --- FUNCTION: PROTECTION OF CELLS FROM COMPLEMENT-MEDIATED DAMAGE (BY
 CC SIMILARITY).
 CC --- SUBCELLULAR LOCATION: Attached to the membrane by a GPI-anchor
 CC --- TISSUE SPECIFICITY: BRAIN, SPLEEN, EPITHELIA, SKELETAL MUSCLE,
 CC LIVER, TESTES, THYMUS, SPLEEN AND LYMPH NODE.
 CC --- DOMAIN: THE FIRST SUSHI DOMAIN (S381) IS NOT NECESSARY FOR

cc POLYLACTAMINE THAT ARE FOUND AT THE NONREDUCING TERMINI OF
cc GLYCOPOLIPIDS)
cc 1 SUBCELLULAR LOCATION: Type I membrane protein.
cc 1 SIMILARITY: TO OTHER SELECTINS/LECTINS
cc 1 SIMILARITY: CONTAINS 1 C-TYPE LECTIN FAMILY DOMAIN.
cc 1 SIMILARITY: CONTAINS 1 EGF-LIKE DOMAIN.
cc 1 SIMILARITY: CONTAINS 5 SUSHI (SCR) DOMAINS.
cc
cc THIS SWISS PROT entry is copyright. It is produced through a collaboration
cc between the Swiss Institute of Bioinformatics and the EMBL Outstation -
cc the European Bioinformatics Institute. There are no restrictions on its
cc use by non profit institutions as long as its content is in no way
cc modified and this statement is not removed. Usage by and for commercial
cc entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
cc or send an email to license@isb-sib.ch)
cc
cc EMBL: L25627; AAA41113.1;
cc RSP: P16581; IGA.
cc InterPro: IPR000961; EGF-like.
cc InterPro: IPR000946; SUSHI SCR cnp.
cc InterPro: IPR001404; Lectin_C.
cc Pfam: PF00008; EGF_1.
cc Pfam: PF00059; Lectin_C; 1.
cc Pfam: PF00084; SUSHI; 5.
cc SMART: SM00042; CTP; 5.
cc SMART: SM00044; CLEFT; 1.
cc SMART: SM00101; EGF_1.
cc PROSITE: PS00222; EGF_1; 1.
cc PROSITE: PS01106; EGF_2; 1.
cc PROSITE: PS06615; C-TYPE LECTIN_1; 1.
cc PROSITE: PS56041; C-TYPE LECTIN_2; 1.
cc
cc Cell adhesion; Transmembrane; Glycoprotein; EGF-like domain; Lectin;
cc Selectin; Signal; SUSHI; Repeat.
cc
cc SIGNAL 1 21 POTENTIAL.
cc CHAIN 22 549 E-SELECTIN.
cc DOMAIN 22 494 EXTRACELLULAR (POTENTIAL).
cc TRANSMEM 495 516 POTENTIAL.
cc DOMAIN 517 549 CYTOPLASMIC (POTENTIAL).
cc
cc DOMAIN 38 148 EGF-LIKE.
cc DOMAIN 169 175 SUSHI 1.
cc DOMAIN 179 249 SUSHI 2.
cc DOMAIN 242 401 SUSHI 3.
cc DOMAIN 404 464 SUSHI 4.
cc DOMAIN 467 427 SUSHI 5.
cc DOMAIN 430 486 SUSHI 6.
cc DISULFID 40 148 HY SIMILARITY.
cc DISULFID 111 140 HY SIMILARITY.
cc DISULFID 143 154 HY SIMILARITY.
cc DISULFID 148 163 HY SIMILARITY.
cc DISULFID 165 174 HY SIMILARITY.
cc DISULFID 180 225 HY SIMILARITY.
cc DISULFID 210 248 HY SIMILARITY.
cc DISULFID 243 287 HY SIMILARITY.
cc DISULFID 273 400 HY SIMILARITY.
cc DISULFID 405 450 HY SIMILARITY.
cc DISULFID 436 463 HY SIMILARITY.
cc DISULFID 468 413 HY SIMILARITY.
cc DISULFID 499 426 HY SIMILARITY.
cc DISULFID 431 472 HY SIMILARITY.
cc DISULFID 458 485 HY SIMILARITY.
cc CARBOHYD 25 25 N-LINKED (GLCNAC...) (POTENTIAL).
cc CARBOHYD 60 60 N-LINKED (GLCNAC...) (POTENTIAL).
cc CARBOHYD 145 145 N-LINKED (GLCNAC...) (POTENTIAL).
cc CARBOHYD 192 192 N-LINKED (GLCNAC...) (POTENTIAL).
cc CARBOHYD 203 203 N-LINKED (GLCNAC...) (POTENTIAL).
cc CARBOHYD 266 266 N-LINKED (GLCNAC...) (POTENTIAL).
cc CARBOHYD 413 413 N-LINKED (GLCNAC...) (POTENTIAL).
cc CARBOHYD 420 320 N-LINKED (GLCNAC...) (POTENTIAL).
cc CARBOHYD 433 443 N-LINKED (GLCNAC...) (POTENTIAL).
cc CARBOHYD 441 441 N-LINKED (GLCNAC...) (POTENTIAL).
cc CARBOHYD 445 465 N-LINKED (GLCNAC...) (POTENTIAL).
cc SEQUENCE 549 AA; 60074 MW; 85CFECDB7B0144CB CRC64;

Query Match 18.6%; Score 140; DH 1; Length 549;
Best local Similarity 29.6%; Pred. No 1.2e-06;
Matches 34; Conservative 13; Mismatches 58; Indels 10; Gaps 4;
QY 16 PYSLDIPVGVLRVTCSPSTRIGERKALFCISNQVIAIWDKAPPCESVNTTISGDP 76
DB 583 PSASAPFGSGSKPKSCDFGELKGRRLQCSQCE... WDSKPKFCAGV... QCSSE 444
QY 76 IYVGGTMAKSGKAPPHGSDVTEFCANFTMKSSKTVWQANEMWGPITALPVTES 140
DB 435 DLFGK MAMSUSGPAVFGIVTEFTEPEGWILNLSILLIARAGWS AMLPTGEA 487
RESULT 24
LEW2 HUMAN
ID LEW2 HUMAN STANDARD; PRT; 510 AA.
AC P.6581; P16581;
DT 01-APR-1993 (Ref. 14, Created)
DT 01-AUG-1999 (Ref. 15, last sequence update)
DT 16-OCT-2001 (Ref. 40, last annotation update)
DE E-selectin precursor (Endothelial leukocyte adhesion molecule 1)
DE (ELAM-1) (Leukocyte-endothelial cell adhesion molecule 2) (LELAM2)
DE (CD62E).
GN SLE OR ELAM.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Gracilata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN 11
RP SEQUENCE FROM N.A.
RX MEDLINE-90175359; PubMed-1689848;
RA Hession C, Osborn L, Goff B, Chappas C, Vassallo C,
RA Paisek M, Pittack C, Tizard R, Gwall S, McCarthy K, Hopple S,
RA Lobb R;
RT "Endothelial leukocyte adhesion molecule 1: direct expression cloning
RT and functional interactions.";
RL Proc. Natl. Acad. Sci. U.S.A. 87:1677-1677(1990).
RN 12
RP SEQUENCE FROM N.A.
RX MEDLINE-86162047; PubMed 2466335;
RA Bevilacqua M.P., Steingell S., Gimbene M.A. Jr., Seed B;
RT "Endothelial leukocyte adhesion molecule 1: an inducible receptor for
RT neutrophils related to complement regulatory proteins and lectins.";
RL Science 243:1150-1155(1989).
RN 13
RP SEQUENCE FROM N.A.
RX MEDLINE-91115870; PubMed-1703529;
RA Collins T, Williams A, Johnston G, Kim J, Eddy R, Shows L;
RA Gimbene M.A. Jr., Bevilacqua M.P.;
RT "Structure and chromosomal location of the gene for endothelial-
RT leukocyte adhesion molecule 1.";
RL J. Biol. Chem. 266:2466-2473(1991).
RN 14
RP LIGAND.
RX MEDLINE-9 068005; PubMed-1701274;
RA Phillips M.L., Nudelma E., Gacta F.C., Perez M., Singhal A.K.,
RA Iakomori S., Paulson J.C.;
RT "ELAM 1 mediates cell adhesion by recognition of a carbohydrate
RT ligand, sialyl-Lex.";
RL Science 240:1130-1132(1990).
RN 15
RP 3D-STRUCTURE MODELING OF LECTIN DOMAIN.
RX MEDLINE-94262575; PubMed-7681016;
RA Mills A.;
RT "Modelling the carbohydrate recognition domain of human E-selectin.";
RL FEBS Lett 319:5-11(1993).
RN 16
RP X-RAY CRYSTALLOGRAPHY (2.0 ANGSTROMS) OF 22 178.
RX MEDLINE-94130646; PubMed-7509040;
RA Graves B.J., Crowther R.L., Chandra S., Rumberger J.M., Li S.,
RA Huang K-S., Presky D.H., Familletti P.G., Welitzky R.A., Burns D.K.;
RT "Insight into E-selectin/ligand interaction from the crystal
RT structure and mutagenesis of the lectin domains.";

81 Nature 367:532-536(1994).
 82 [1]
 83 VARIANT ARG 149.
 84 MEDLINE-95179107; PubMed 7546025;
 85 Wondol K., Felix S., Klier F.X., Bruchard R., Monke T., Sebhatke S.,
 86 Schuler K., Glaser C., Rohde K., Baumann G., Speer A.:
 87 "p-selectin polymorphism and atherosclerosis: an association study.";
 88 Hum. Mol. Genet. 3:1905-1913(1994).
 89 [8]
 90 VARIANT ARG 149.
 91 MEDLINE-99144008; PubMed 9903708;
 92 Ye S.Q., Ishii D., Virsik D., Zhang L.Q., Yoshim S.E., Gupta R.:
 93 "A p-selectin polymorphism affects the mutation of serine 128 to arginine in
 94 the EGF gene: a risk factor for coronary artery disease.";
 95 J. Biomed. Sci. 6:18-21(1999).
 96 [1]
 97 FUNCTION: EXPRESSION CYTOKINE-INDUCED ENDOTHELIAL CELLS AND
 98 MEDIATES THEIR BINDING TO LEUKOCYTES. THE LIGAND RECOGNIZED BY
 99 PLAM-1 IS STALKED LECTIN X (ALPHA(1-3)FUCOSYLATED DERIVATIVES OF
 100 POLYACETAMINE THAT ARE FOUND AT THE NONREDUCING TERMINI OF
 101 GLYCOPROTEINS).
 102 [1]
 103 SUBCELLULAR LOCATION: Type I membrane protein.
 104 [1]
 105 POLYMORPHISM: A POLYMORPHISM IN POSITION 149 IS ASSOCIATED WITH A
 106 HIGHER RISK OF CORONARY ARTERY DISEASE (CAD). A SIGNIFICANTLY
 107 HIGHER MUTATION FREQUENCY (ARG 149) IS OBSERVED IN PATIENTS WITH
 108 ANGIOGRAPHICALLY PROVEN SEVERE ATHEROSCLEROSIS COMPARED WITH AN
 109 UNSELECTED POPULATION (SER 149).
 110 [1]
 111 SIMILARITY: TO OTHER SELECTINS/LECTINS.
 112 [1]
 113 SIMILARITY: CONTAINS 1 C- TYPE LECTIN FAMILY DOMAIN.
 114 [1]
 115 SIMILARITY: CONTAINS 1 EGF-LIKE DOMAIN.
 116 [1]
 117 SIMILARITY: CONTAINS 6 SUSHI (SCR) DOMAINS.
 118 [1]
 119 DATABASE NAME: PROTEIN DATA BANK (PDB) ENTRY: 1L21;
 120 WWW: <http://www.ncbi.nlm.nih.gov/Structure/1L21>.
 121 [1]
 122 THIS SWISS-PROT entry is copyrighted. It is produced through a collaboration
 123 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 124 the European Bioinformatics Institute. There are no restrictions on its
 125 use by non-profit institutions as long as its content is in no way
 126 modified and this statement is not removed. Usage by and for commercial
 127 entities requires a license agreement (<http://www.isdb.bham.ac.uk>)
 128 or send an email to license@isb.stb.ch.
 129 [1]
 130 EMBL: M30307; AAA^2377; 1;
 131 EMBL: M61894; AAA^2378; 1; JOINED.
 132 EMBL: M61895; AAA^2378; 1; JOINED.
 133 EMBL: M61896; AAA^2378; 1; JOINED.
 134 EMBL: M61897; AAA^2378; 1; JOINED.
 135 EMBL: M61898; AAA^2378; 1; JOINED.
 136 EMBL: M61899; AAA^2378; 1; JOINED.
 137 EMBL: M61900; AAA^2378; 1; JOINED.
 138 EMBL: M61901; AAA^2378; 1; JOINED.
 139 EMBL: M61902; AAA^2378; 1; JOINED.
 140 EMBL: M61903; AAA^2378; 1; JOINED.
 141 EMBL: M61904; AAA^2378; 1; JOINED.
 142 EMBL: M61905; AAA^2378; 1; JOINED.
 143 EMBL: M61906; AAA^2378; 1; JOINED.
 144 EMBL: M61907; AAA^2378; 1; JOINED.
 145 EMBL: M61908; AAA^2378; 1; JOINED.
 146 EMBL: M61909; AAA^2378; 1; JOINED.
 147 EMBL: M61910; AAA^2378; 1; JOINED.
 148 EMBL: M61911; AAA^2378; 1; JOINED.
 149 EMBL: M61912; AAA^2378; 1; JOINED.
 150 EMBL: M61913; AAA^2378; 1; JOINED.
 151 EMBL: M61914; AAA^2378; 1; JOINED.
 152 EMBL: M61915; AAA^2378; 1; JOINED.
 153 EMBL: M61916; AAA^2378; 1; JOINED.
 154 EMBL: M61917; AAA^2378; 1; JOINED.
 155 EMBL: M61918; AAA^2378; 1; JOINED.
 156 EMBL: M61919; AAA^2378; 1; JOINED.
 157 EMBL: M61920; AAA^2378; 1; JOINED.
 158 EMBL: M61921; AAA^2378; 1; JOINED.
 159 EMBL: M61922; AAA^2378; 1; JOINED.
 160 EMBL: M61923; AAA^2378; 1; JOINED.
 161 EMBL: M61924; AAA^2378; 1; JOINED.
 162 EMBL: M61925; AAA^2378; 1; JOINED.
 163 EMBL: M61926; AAA^2378; 1; JOINED.
 164 EMBL: M61927; AAA^2378; 1; JOINED.
 165 EMBL: M61928; AAA^2378; 1; JOINED.
 166 EMBL: M61929; AAA^2378; 1; JOINED.
 167 EMBL: M61930; AAA^2378; 1; JOINED.
 168 EMBL: M61931; AAA^2378; 1; JOINED.
 169 EMBL: M61932; AAA^2378; 1; JOINED.
 170 EMBL: M61933; AAA^2378; 1; JOINED.
 171 EMBL: M61934; AAA^2378; 1; JOINED.
 172 EMBL: M61935; AAA^2378; 1; JOINED.
 173 EMBL: M61936; AAA^2378; 1; JOINED.
 174 EMBL: M61937; AAA^2378; 1; JOINED.
 175 EMBL: M61938; AAA^2378; 1; JOINED.
 176 EMBL: M61939; AAA^2378; 1; JOINED.
 177 EMBL: M61940; AAA^2378; 1; JOINED.
 178 EMBL: M61941; AAA^2378; 1; JOINED.
 179 EMBL: M61942; AAA^2378; 1; JOINED.
 180 EMBL: M61943; AAA^2378; 1; JOINED.
 181 EMBL: M61944; AAA^2378; 1; JOINED.
 182 EMBL: M61945; AAA^2378; 1; JOINED.
 183 EMBL: M61946; AAA^2378; 1; JOINED.
 184 EMBL: M61947; AAA^2378; 1; JOINED.
 185 EMBL: M61948; AAA^2378; 1; JOINED.
 186 EMBL: M61949; AAA^2378; 1; JOINED.
 187 EMBL: M61950; AAA^2378; 1; JOINED.
 188 EMBL: M61951; AAA^2378; 1; JOINED.
 189 EMBL: M61952; AAA^2378; 1; JOINED.
 190 EMBL: M61953; AAA^2378; 1; JOINED.
 191 EMBL: M61954; AAA^2378; 1; JOINED.
 192 EMBL: M61955; AAA^2378; 1; JOINED.
 193 EMBL: M61956; AAA^2378; 1; JOINED.
 194 EMBL: M61957; AAA^2378; 1; JOINED.
 195 EMBL: M61958; AAA^2378; 1; JOINED.
 196 EMBL: M61959; AAA^2378; 1; JOINED.
 197 EMBL: M61960; AAA^2378; 1; JOINED.
 198 EMBL: M61961; AAA^2378; 1; JOINED.
 199 EMBL: M61962; AAA^2378; 1; JOINED.
 200 EMBL: M61963; AAA^2378; 1; JOINED.
 201 EMBL: M61964; AAA^2378; 1; JOINED.
 202 EMBL: M61965; AAA^2378; 1; JOINED.
 203 EMBL: M61966; AAA^2378; 1; JOINED.
 204 EMBL: M61967; AAA^2378; 1; JOINED.
 205 EMBL: M61968; AAA^2378; 1; JOINED.
 206 EMBL: M61969; AAA^2378; 1; JOINED.
 207 EMBL: M61970; AAA^2378; 1; JOINED.
 208 EMBL: M61971; AAA^2378; 1; JOINED.
 209 EMBL: M61972; AAA^2378; 1; JOINED.
 210 EMBL: M61973; AAA^2378; 1; JOINED.
 211 EMBL: M61974; AAA^2378; 1; JOINED.
 212 EMBL: M61975; AAA^2378; 1; JOINED.
 213 EMBL: M61976; AAA^2378; 1; JOINED.
 214 EMBL: M61977; AAA^2378; 1; JOINED.
 215 EMBL: M61978; AAA^2378; 1; JOINED.
 216 EMBL: M61979; AAA^2378; 1; JOINED.
 217 EMBL: M61980; AAA^2378; 1; JOINED.
 218 EMBL: M61981; AAA^2378; 1; JOINED.
 219 EMBL: M61982; AAA^2378; 1; JOINED.
 220 EMBL: M61983; AAA^2378; 1; JOINED.
 221 EMBL: M61984; AAA^2378; 1; JOINED.
 222 EMBL: M61985; AAA^2378; 1; JOINED.
 223 EMBL: M61986; AAA^2378; 1; JOINED.
 224 EMBL: M61987; AAA^2378; 1; JOINED.
 225 EMBL: M61988; AAA^2378; 1; JOINED.
 226 EMBL: M61989; AAA^2378; 1; JOINED.
 227 EMBL: M61990; AAA^2378; 1; JOINED.
 228 EMBL: M61991; AAA^2378; 1; JOINED.
 229 EMBL: M61992; AAA^2378; 1; JOINED.
 230 EMBL: M61993; AAA^2378; 1; JOINED.
 231 EMBL: M61994; AAA^2378; 1; JOINED.
 232 EMBL: M61995; AAA^2378; 1; JOINED.
 233 EMBL: M61996; AAA^2378; 1; JOINED.
 234 EMBL: M61997; AAA^2378; 1; JOINED.
 235 EMBL: M61998; AAA^2378; 1; JOINED.
 236 EMBL: M61999; AAA^2378; 1; JOINED.
 237 EMBL: M62000; AAA^2378; 1; JOINED.
 238 EMBL: M62001; AAA^2378; 1; JOINED.
 239 EMBL: M62002; AAA^2378; 1; JOINED.
 240 EMBL: M62003; AAA^2378; 1; JOINED.
 241 EMBL: M62004; AAA^2378; 1; JOINED.
 242 EMBL: M62005; AAA^2378; 1; JOINED.
 243 EMBL: M62006; AAA^2378; 1; JOINED.
 244 EMBL: M62007; AAA^2378; 1; JOINED.
 245 EMBL: M62008; AAA^2378; 1; JOINED.
 246 EMBL: M62009; AAA^2378; 1; JOINED.
 247 EMBL: M62010; AAA^2378; 1; JOINED.
 248 EMBL: M62011; AAA^2378; 1; JOINED.
 249 EMBL: M62012; AAA^2378; 1; JOINED.
 250 EMBL: M62013; AAA^2378; 1; JOINED.
 251 EMBL: M62014; AAA^2378; 1; JOINED.
 252 EMBL: M62015; AAA^2378; 1; JOINED.
 253 EMBL: M62016; AAA^2378; 1; JOINED.
 254 EMBL: M62017; AAA^2378; 1; JOINED.
 255 EMBL: M62018; AAA^2378; 1; JOINED.
 256 EMBL: M62019; AAA^2378; 1; JOINED.
 257 EMBL: M62020; AAA^2378; 1; JOINED.
 258 EMBL: M62021; AAA^2378; 1; JOINED.
 259 EMBL: M62022; AAA^2378; 1; JOINED.
 260 EMBL: M62023; AAA^2378; 1; JOINED.
 261 EMBL: M62024; AAA^2378; 1; JOINED.
 262 EMBL: M62025; AAA^2378; 1; JOINED.
 263 EMBL: M62026; AAA^2378; 1; JOINED.
 264 EMBL: M62027; AAA^2378; 1; JOINED.
 265 EMBL: M62028; AAA^2378; 1; JOINED.
 266 EMBL: M62029; AAA^2378; 1; JOINED.
 267 EMBL: M62030; AAA^2378; 1; JOINED.
 268 EMBL: M62031; AAA^2378; 1; JOINED.
 269 EMBL: M62032; AAA^2378; 1; JOINED.
 270 EMBL: M62033; AAA^2378; 1; JOINED.
 271 EMBL: M62034; AAA^2378; 1; JOINED.
 272 EMBL: M62035; AAA^2378; 1; JOINED.
 273 EMBL: M62036; AAA^2378; 1; JOINED.
 274 EMBL: M62037; AAA^2378; 1; JOINED.
 275 EMBL: M62038; AAA^2378; 1; JOINED.
 276 EMBL: M62039; AAA^2378; 1; JOINED.
 277 EMBL: M62040; AAA^2378; 1; JOINED.
 278 EMBL: M62041; AAA^2378; 1; JOINED.
 279 EMBL: M62042; AAA^2378; 1; JOINED.
 280 EMBL: M62043; AAA^2378; 1; JOINED.
 281 EMBL: M62044; AAA^2378; 1; JOINED.
 282 EMBL: M62045; AAA^2378; 1; JOINED.
 283 EMBL: M62046; AAA^2378; 1; JOINED.
 284 EMBL: M62047; AAA^2378; 1; JOINED.
 285 EMBL: M62048; AAA^2378; 1; JOINED.
 286 EMBL: M62049; AAA^2378; 1; JOINED.
 287 EMBL: M62050; AAA^2378; 1; JOINED.
 288 EMBL: M62051; AAA^2378; 1; JOINED.
 289 EMBL: M62052; AAA^2378; 1; JOINED.
 290 EMBL: M62053; AAA^2378; 1; JOINED.
 291 EMBL: M62054; AAA^2378; 1; JOINED.
 292 EMBL: M62055; AAA^2378; 1; JOINED.
 293 EMBL: M62056; AAA^2378; 1; JOINED.
 294 EMBL: M62057; AAA^2378; 1; JOINED.
 295 EMBL: M62058; AAA^2378; 1; JOINED.
 296 EMBL: M62059; AAA^2378; 1; JOINED.
 297 EMBL: M62060; AAA^2378; 1; JOINED.
 298 EMBL: M62061; AAA^2378; 1; JOINED.
 299 EMBL: M62062; AAA^2378; 1; JOINED.
 300 EMBL: M62063; AAA^2378; 1; JOINED.
 301 EMBL: M62064; AAA^2378; 1; JOINED.
 302 EMBL: M62065; AAA^2378; 1; JOINED.
 303 EMBL: M62066; AAA^2378; 1; JOINED.
 304 EMBL: M62067; AAA^2378; 1; JOINED.
 305 EMBL: M62068; AAA^2378; 1; JOINED.
 306 EMBL: M62069; AAA^2378; 1; JOINED.
 307 EMBL: M62070; AAA^2378; 1; JOINED.
 308 EMBL: M62071; AAA^2378; 1; JOINED.
 309 EMBL: M62072; AAA^2378; 1; JOINED.
 310 EMBL: M62073; AAA^2378; 1; JOINED.
 311 EMBL: M62074; AAA^2378; 1; JOINED.
 312 EMBL: M62075; AAA^2378; 1; JOINED.
 313 EMBL: M62076; AAA^2378; 1; JOINED.
 314 EMBL: M62077; AAA^2378; 1; JOINED.
 315 EMBL: M62078; AAA^2378; 1; JOINED.
 316 EMBL: M62079; AAA^2378; 1; JOINED.
 317 EMBL: M62080; AAA^2378; 1; JOINED.
 318 EMBL: M62081; AAA^2378; 1; JOINED.
 319 EMBL: M62082; AAA^2378; 1; JOINED.
 320 EMBL: M62083; AAA^2378; 1; JOINED.
 321 EMBL: M62084; AAA^2378; 1; JOINED.
 322 EMBL: M62085; AAA^2378; 1; JOINED.
 323 EMBL: M62086; AAA^2378; 1; JOINED.
 324 EMBL: M62087; AAA^2378; 1; JOINED.
 325 EMBL: M62088; AAA^2378; 1; JOINED.
 326 EMBL: M62089; AAA^2378; 1; JOINED.
 327 EMBL: M62090; AAA^2378; 1; JOINED.
 328 EMBL: M62091; AAA^2378; 1; JOINED.
 329 EMBL: M62092; AAA^2378; 1; JOINED.
 330 EMBL: M62093; AAA^2378; 1; JOINED.
 331 EMBL: M62094; AAA^2378; 1; JOINED.
 332 EMBL: M62095; AAA^2378; 1; JOINED.
 333 EMBL: M62096; AAA^2378; 1; JOINED.
 334 EMBL: M62097; AAA^2378; 1; JOINED.
 335 EMBL: M62098; AAA^2378; 1; JOINED.
 336 EMBL: M62099; AAA^2378; 1; JOINED.
 337 EMBL: M62100; AAA^2378; 1; JOINED.
 338 EMBL: M62101; AAA^2378; 1; JOINED.
 339 EMBL: M62102; AAA^2378; 1; JOINED.
 340 EMBL: M62103; AAA^2378; 1; JOINED.
 341 EMBL: M62104; AAA^2378; 1; JOINED.
 342 EMBL: M62105; AAA^2378; 1; JOINED.
 343 EMBL: M62106; AAA^2378; 1; JOINED.
 344 EMBL: M62107; AAA^2378; 1; JOINED.
 345 EMBL: M62108; AAA^2378; 1; JOINED.
 346 EMBL: M62109; AAA^2378; 1; JOINED.
 347 EMBL: M62110; AAA^2378; 1; JOINED.
 348 EMBL: M62111; AAA^2378; 1; JOINED.
 349 EMBL: M62112; AAA^2378; 1; JOINED.
 350 EMBL: M62113; AAA^2378; 1; JOINED.
 351 EMBL: M62114; AAA^2378; 1; JOINED.
 352 EMBL: M62115; AAA^2378; 1; JOINED.
 353 EMBL: M62116; AAA^2378; 1; JOINED.
 354 EMBL: M62117; AAA^2378; 1; JOINED.
 355 EMBL: M62118; AAA^2378; 1; JOINED.
 356 EMBL: M62119; AAA^2378; 1; JOINED.
 357 EMBL: M62120; AAA^2378; 1; JOINED.
 358 EMBL: M62121; AAA^2378; 1; JOINED.
 359 EMBL: M62122; AAA^2378; 1; JOINED.
 360 EMBL: M62123; AAA^2378; 1; JOINED.
 361 EMBL: M62124; AAA^2378; 1; JOINED.
 362 EMBL: M62125; AAA^2378; 1; JOINED.
 363 EMBL: M62126; AAA^2378; 1; JOINED.
 364 EMBL: M62127; AAA^2378; 1; JOINED.
 365 EMBL: M62128; AAA^2378; 1; JOINED.
 366 EMBL: M62129; AAA^2378; 1; JOINED.
 367 EMBL: M62130; AAA^2378; 1; JOINED.
 368 EMBL: M62131; AAA^2378; 1; JOINED.
 369 EMBL: M62132; AAA^2378; 1; JOINED.
 370 EMBL: M62133; AAA^2378; 1; JOINED.
 371 EMBL: M62134; AAA^2378; 1; JOINED.
 372 EMBL: M62135; AAA^2378; 1; JOINED.
 373 EMBL: M62136; AAA^2378; 1; JOINED.
 374 EMBL: M62137; AAA^2378; 1; JOINED.
 375 EMBL: M62138; AAA^2378; 1; JOINED.
 376 EMBL: M62139; AAA^2378; 1; JOINED.
 377 EMBL: M62140; AAA^2378; 1; JOINED.
 378 EMBL: M62141; AAA^2378; 1; JOINED.
 379 EMBL: M62142; AAA^2378; 1; JOINED.
 380 EMBL: M62143; AAA^2378; 1; JOINED.
 381 EMBL: M62144; AAA^2378; 1; JOINED.
 382 EMBL: M62145; AAA^2378; 1; JOINED.
 383 EMBL: M62146; AAA^2378; 1; JOINED.
 384 EMBL: M62147; AAA^2378; 1; JOINED.
 385 EMBL: M62148; AAA^2378; 1; JOINED.
 386 EMBL: M62149; AAA^2378; 1; JOINED.
 387 EMBL: M62150; AAA^2378; 1; JOINED.
 388 EMBL: M62151; AAA^2378; 1; JOINED.
 389 EMBL: M62152; AAA^2378; 1; JOINED.
 390 EMBL: M62153; AAA^2378; 1; JOINED.
 391 EMBL: M62154; AAA^2378; 1; JOINED.
 392 EMBL: M62155; AAA^2378; 1; JOINED.
 393 EMBL: M62156; AAA^2378; 1; JOINED.
 394 EMBL: M62157; AAA^2378; 1; JOINED.
 395 EMBL: M62158; AAA^2378; 1; JOINED.
 396 EMBL: M62159; AAA^2378; 1; JOINED.
 397 EMBL: M62160; AAA^2378; 1; JOINED.
 398 EMBL: M62161; AAA^2378; 1; JOINED.
 399 EMBL: M62162; AAA^2378; 1; JOINED.
 400 EMBL: M62163; AAA^2378; 1; JOINED.
 401 EMBL: M62164; AAA^2378; 1; JOINED.
 402 EMBL: M62165; AAA^2378; 1; JOINED.
 403 EMBL: M62166; AAA^2378; 1; JOINED.
 404 EMBL: M62167; AAA^2378; 1; JOINED.
 405 EMBL: M62168; AAA^2378; 1; JOINED.
 406 EMBL: M62169; AAA^2378; 1; JOINED.
 407 EMBL: M62170; AAA^2378; 1; JOINED.
 408 EMBL: M62171; AAA^2378; 1; JOINED.
 409 EMBL: M62172; AAA^2378; 1; JOINED.
 410 EMBL: M62173; AAA^2378; 1; JOINED.
 411 EMBL: M62174; AAA^2378; 1; JOINED.
 412 EMBL: M62175; AAA^2378; 1; JOINED.
 413 EMBL: M62176; AAA^2378; 1; JOINED.
 414 EMBL: M62177; AAA^2378; 1; JOINED.
 415 EMBL: M62178; AAA^2378; 1; JOINED.
 416 EMBL: M62179; AAA^2378; 1; JOINED.
 417 EMBL: M62180; AAA^2378; 1; JOINED.
 418 EMBL: M62181; AAA^2378; 1; JOINED.
 419 EMBL: M62182; AAA^2378; 1; JOINED.
 420 EMBL: M62183; AAA^2378; 1; JOINED.
 421 EMBL: M62184; AAA^2378; 1; JOINED.
 422 EMBL: M62185; AAA^2378; 1; JOINED.
 423 EMBL: M62186; AAA^2378; 1; JOINED.
 424 EMBL: M62187; AAA^2378; 1; JOINED.
 425 EMBL: M62188; AAA^2378; 1; JOINED.
 426 EMBL: M62189; AAA^2378; 1; JOINED.
 427 EMBL: M62190; AAA^2378; 1; JOINED.
 428 EMBL: M62191; AAA^2378; 1; JOINED.
 429 EMBL: M62192; AAA^2378; 1; JOINED.
 430 EMBL: M62193; AAA^2378; 1; JOINED.
 431 EMBL: M62194; AAA^2378; 1; JOINED.
 432 EMBL: M62195; AAA^2378; 1; JOINED.
 433 EMBL: M62196; AAA^2378; 1; JOINED.
 434 EMBL: M62197; AAA^2378; 1; JOINED.
 435 EMBL: M62198; AAA^2378; 1; JOINED.
 436 EMBL: M62199; AAA^2378; 1; JOINED.
 437 EMBL: M62200; AAA^2378; 1; JOINED.
 438 EMBL: M62201; AAA^2378; 1; JOINED.
 439 EMBL: M62202; AAA^2378; 1; JOINED.
 440 EMBL: M62203; AAA^2378; 1; JOINED.
 441 EMBL: M62204; AAA^2378; 1; JOINED.
 442 EMBL: M62205; AAA^2378; 1; JOINED.
 443 EMBL: M62206; AAA^2378; 1; JOINED.
 444 EMBL: M62207; AAA^2378; 1; JOINED.
 445 EMBL: M62208; AAA^2378; 1; JOINED.
 446 EMBL: M62209; AAA^2378; 1; JOINED.
 447 EMBL: M62210; AAA^2378; 1; JOINED.
 448 EMBL: M62211; AAA^2378; 1; JOINED.
 449 EMBL: M62212; AAA^2378; 1; JOINED.
 450 EMBL: M62213; AAA^2378; 1; JOINED.
 451 EMBL: M62214; AAA^2378; 1; JOINED.
 452 EMBL: M62215; AAA^2378; 1; JOINED.
 453 EMBL: M62216; AAA^2378; 1; JOINED.
 454 EMBL: M62217; AAA^2378; 1; JOINED.
 455 EMBL: M62218; AAA^2378; 1; JOINED.
 456 EMBL: M62219; AAA^2378; 1; JOINED.
 457 EMBL: M62220; AAA^2378; 1; JOINED.
 458 EMBL: M62221; AAA^2378; 1; JOINED.
 459 EMBL: M62222; AAA^2378; 1; JOINED.
 460 EMBL: M62223; AAA^2378; 1; JOINED.
 461 EMBL: M62224; AAA^2378; 1; JOINED.
 462 EMBL: M62225; AAA^2378; 1; JOINED.
 463 EMBL: M62226; AAA^2378; 1; JOINED.
 464 EMBL: M62227; AAA^2378; 1; JOINED.
 465 EMBL: M62228; AAA^2378; 1; JOINED.
 466 EMBL: M62229; AAA^2378; 1; JOINED.
 467 EMBL: M62230; AAA^2378; 1; JOINED.
 468 EMBL: M62231; AAA^2378; 1; JOINED.
 469 EMBL: M62232; AAA^2378; 1; JOINED.
 470 EMBL: M62233; AAA^2378; 1; JOINED.
 471 EMBL: M62234; AAA^2378; 1; JOINED.
 472 EMBL: M62235; AAA^2378; 1; JOINED.
 473 EMBL: M62236; AAA^2378; 1; JOINED.
 474 EMBL: M62237; AAA^2378; 1; JOINED.
 475 EMBL: M62238; AAA^2378; 1; JOINED.
 476 EMBL: M62239; AAA^2378; 1; JOINED.
 477 EMBL: M62240; AAA^2378; 1; JOINED.
 478 EMBL: M62241; AAA^2378; 1; JOINED.
 479 EMBL: M62242; AAA^2378; 1; JOINED.
 480 EMBL: M62243; AAA^2378; 1; JOINED.
 481 EMBL: M62244; AAA^2378; 1; JOINED.
 482 EMBL: M62245; AAA^2378; 1; JOINED.
 483 EMBL: M62246; AAA^2378; 1; JOINED.
 484 EMBL: M62247; AAA^2378; 1; JOINED.
 485 EMBL: M62248; AAA^2378; 1; JOINED.
 486 EMBL: M62249; AAA^2378; 1; JOINED.
 487 EMBL: M62250; AAA^2378; 1; JOINED.
 488 EMBL:

DT 14 AUG-1987 (rel. 05, Created)
 DT 01 AUG-1990 (rel. 15, Last sequence update)
 DI 16 OCT-2001 (rel. 40, Last annotation update)
 DE Coagulation factor XIII B chain precursor (EC 2.3.2.13) (protein-
 DE glutamine gamma-glutamyltransferase B chain) (Transglutaminase B
 DE chain).
 GN F13B.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=91105154; PubMed=2271707;
 RA Bottemus R.E., Ichinose A., Davie E.W.;
 RT "Nucleotide sequence of the gene for the b subunit of human factor
 RT XIII.";
 RL Biochemistry 29:11195-11209(1990).
 [2]
 RP SEQUENCE OF 2 661 FROM N.A.
 RX MEDLINE=87026535; PubMed=4021194;
 RA Ichinose A., McMahon H.A., Fujikawa K., Davie E.W.;
 RT "Amino acid sequence of the b subunit of human factor XIII, a protein
 RT composed of ten repetitive segments.";
 RL Biochemistry 25:4633-4638(1986).
 [3]
 RP REVISIONS.
 RA Ichinose A.;
 RL Submitted (FEB-1987) to the EMBL/GenBank/DBJ databases.
 [4]
 RP SEQUENCE OF 1-20 FROM N.A.
 RX MEDLINE=90251467; PubMed=2339367;
 RA Grundmann U., Norlich C., Rein T., Zettlmeissl G.;
 RT "Complete cDNA sequence encoding the b subunit of human factor XIII.";
 RL Nucleic Acids Res. 18:2817-2817(1990).
 [5]
 RP VARIANT PHE 450.
 RX MEDLINE=94414189; PubMed=8424218;
 RA Bashiguchi T., Saito M., Morishita E., Matsuda T., Ichinose A.;
 RT "Two genetic defects in a patient with complete deficiency of the b-
 RT subunit for coagulation factor XIII.";
 RL Blood 82:145-150(1993).
 [6]
 RP FUNCTION: THE B CHAIN OF FACTOR XIII IS NOT CATALYTICALLY ACTIVE,
 RP BUT IS THOUGHT TO STABILIZE THE A SUBUNITS AND REGULATE THE RATE
 RP OF TRANSGLUTAMINASE FORMATION BY THROMBIN.
 [7]
 RP SUBUNIT: TETRAMER OF TWO A CHAINS AND TWO B CHAINS.
 [8]
 RP DISEASE: A DEFICIENCY IN F13B CAN RESULT IN A LIFE-LONG BLEEDING
 RP TENDENCY, DEFECTIVE WOUND HEALING, AND HABITUAL ABORTION.
 [9]
 RP SIMILARITY: CONTAINS 10 SUSHI (SCR) DOMAINS.
 [10]
 RP THIS SWISS-PROT entry is copyright. It is produced through a collaboration
 RP between the Swiss Institute of Bioinformatics and the EMBL outstation -
 RP the European Bioinformatics Institute. There are no restrictions on its
 RP use by non-profit institutions as long as its content is in no way
 RP modified and this statement is not removed, usage by and for commercial
 RP entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 RP or send an email to license@isb-sib.ch).
 [11]
 DR EMBL: M64554; AAA51821.1; ALT_SEQ.
 DR EMBL: M14057; AAA88042.1;
 DR EMBL: X51923; CAA56123.1;
 DR PIR: A23840; A23840.
 DR PIR: A36397; A36397.
 DR PIR: S09480; S09480.
 DR ISSP: P06603; IHEI.
 DR MIM: 144580;
 DR InterPro: IPR000436; Sushi_SCR_CCP.
 DR Pfam: PF00084; sushi; 8.
 DR SMART: SM00042; CCP; 8.
 DR Transferrase; Plasma; Blood coagulation; Repeat; Glycoprotein; Signal;
 KW Sushi; Disease mutation.
 FT SIGNAL 1 20

FT CHAIN 21 661
 FT DOMAIN 24 88
 FT SUSHI 1.
 FT SUSHI 2.
 FT SUSHI 3.
 FT SUSHI 4.
 FT SUSHI 5.
 FT SUSHI 6.
 FT SUSHI 7.
 FT SUSHI 8.
 FT SUSHI 9.
 FT SUSHI 10.
 FT SUSHI 11.
 FT SUSHI 12.
 FT SUSHI 13.
 FT SUSHI 14.
 FT SUSHI 15.
 FT SUSHI 16.
 FT SUSHI 17.
 FT SUSHI 18.
 FT SUSHI 19.
 FT SUSHI 20.
 FT SUSHI 21.
 FT SUSHI 22.
 FT SUSHI 23.
 FT SUSHI 24.
 FT SUSHI 25.
 FT SUSHI 26.
 FT SUSHI 27.
 FT SUSHI 28.
 FT SUSHI 29.
 FT SUSHI 30.
 FT SUSHI 31.
 FT SUSHI 32.
 FT SUSHI 33.
 FT SUSHI 34.
 FT SUSHI 35.
 FT SUSHI 36.
 FT SUSHI 37.
 FT SUSHI 38.
 FT SUSHI 39.
 FT SUSHI 40.
 FT SUSHI 41.
 FT SUSHI 42.
 FT SUSHI 43.
 FT SUSHI 44.
 FT SUSHI 45.
 FT SUSHI 46.
 FT SUSHI 47.
 FT SUSHI 48.
 FT SUSHI 49.
 FT SUSHI 50.
 FT SUSHI 51.
 FT SUSHI 52.
 FT SUSHI 53.
 FT SUSHI 54.
 FT SUSHI 55.
 FT SUSHI 56.
 FT SUSHI 57.
 FT SUSHI 58.
 FT SUSHI 59.
 FT SUSHI 60.
 FT SUSHI 61.
 FT SUSHI 62.
 FT SUSHI 63.
 FT SUSHI 64.
 FT SUSHI 65.
 FT SUSHI 66.
 FT SUSHI 67.
 FT SUSHI 68.
 FT SUSHI 69.
 FT SUSHI 70.
 FT SUSHI 71.
 FT SUSHI 72.
 FT SUSHI 73.
 FT SUSHI 74.
 FT SUSHI 75.
 FT SUSHI 76.
 FT SUSHI 77.
 FT SUSHI 78.
 FT SUSHI 79.
 FT SUSHI 80.
 FT SUSHI 81.
 FT SUSHI 82.
 FT SUSHI 83.
 FT SUSHI 84.
 FT SUSHI 85.
 FT SUSHI 86.
 FT SUSHI 87.
 FT SUSHI 88.
 FT SUSHI 89.
 FT SUSHI 90.
 FT SUSHI 91.
 FT SUSHI 92.
 FT SUSHI 93.
 FT SUSHI 94.
 FT SUSHI 95.
 FT SUSHI 96.
 FT SUSHI 97.
 FT SUSHI 98.
 FT SUSHI 99.
 FT SUSHI 100.
 FT SUSHI 101.
 FT SUSHI 102.
 FT SUSHI 103.
 FT SUSHI 104.
 FT SUSHI 105.
 FT SUSHI 106.
 FT SUSHI 107.
 FT SUSHI 108.
 FT SUSHI 109.
 FT SUSHI 110.
 FT SUSHI 111.
 FT SUSHI 112.
 FT SUSHI 113.
 FT SUSHI 114.
 FT SUSHI 115.
 FT SUSHI 116.
 FT SUSHI 117.
 FT SUSHI 118.
 FT SUSHI 119.
 FT SUSHI 120.
 FT SUSHI 121.
 FT SUSHI 122.
 FT SUSHI 123.
 FT SUSHI 124.
 FT SUSHI 125.
 FT SUSHI 126.
 FT SUSHI 127.
 FT SUSHI 128.
 FT SUSHI 129.
 FT SUSHI 130.
 FT SUSHI 131.
 FT SUSHI 132.
 FT SUSHI 133.
 FT SUSHI 134.
 FT SUSHI 135.
 FT SUSHI 136.
 FT SUSHI 137.
 FT SUSHI 138.
 FT SUSHI 139.
 FT SUSHI 140.
 FT SUSHI 141.
 FT SUSHI 142.
 FT SUSHI 143.
 FT SUSHI 144.
 FT SUSHI 145.
 FT SUSHI 146.
 FT SUSHI 147.
 FT SUSHI 148.
 FT SUSHI 149.
 FT SUSHI 150.
 FT SUSHI 151.
 FT SUSHI 152.
 FT SUSHI 153.
 FT SUSHI 154.
 FT SUSHI 155.
 FT SUSHI 156.
 FT SUSHI 157.
 FT SUSHI 158.
 FT SUSHI 159.
 FT SUSHI 160.
 FT SUSHI 161.
 FT SUSHI 162.
 FT SUSHI 163.
 FT SUSHI 164.
 FT SUSHI 165.
 FT SUSHI 166.
 FT SUSHI 167.
 FT SUSHI 168.
 FT SUSHI 169.
 FT SUSHI 170.
 FT SUSHI 171.
 FT SUSHI 172.
 FT SUSHI 173.
 FT SUSHI 174.
 FT SUSHI 175.
 FT SUSHI 176.
 FT SUSHI 177.
 FT SUSHI 178.
 FT SUSHI 179.
 FT SUSHI 180.
 FT SUSHI 181.
 FT SUSHI 182.
 FT SUSHI 183.
 FT SUSHI 184.
 FT SUSHI 185.
 FT SUSHI 186.
 FT SUSHI 187.
 FT SUSHI 188.
 FT SUSHI 189.
 FT SUSHI 190.
 FT SUSHI 191.
 FT SUSHI 192.
 FT SUSHI 193.
 FT SUSHI 194.
 FT SUSHI 195.
 FT SUSHI 196.
 FT SUSHI 197.
 FT SUSHI 198.
 FT SUSHI 199.
 FT SUSHI 200.
 FT SUSHI 201.
 FT SUSHI 202.
 FT SUSHI 203.
 FT SUSHI 204.
 FT SUSHI 205.
 FT SUSHI 206.
 FT SUSHI 207.
 FT SUSHI 208.
 FT SUSHI 209.
 FT SUSHI 210.
 FT SUSHI 211.
 FT SUSHI 212.
 FT SUSHI 213.
 FT SUSHI 214.
 FT SUSHI 215.
 FT SUSHI 216.
 FT SUSHI 217.
 FT SUSHI 218.
 FT SUSHI 219.
 FT SUSHI 220.
 FT SUSHI 221.
 FT SUSHI 222.
 FT SUSHI 223.
 FT SUSHI 224.
 FT SUSHI 225.
 FT SUSHI 226.
 FT SUSHI 227.
 FT SUSHI 228.
 FT SUSHI 229.
 FT SUSHI 230.
 FT SUSHI 231.
 FT SUSHI 232.
 FT SUSHI 233.
 FT SUSHI 234.
 FT SUSHI 235.
 FT SUSHI 236.
 FT SUSHI 237.
 FT SUSHI 238.
 FT SUSHI 239.
 FT SUSHI 240.
 FT SUSHI 241.
 FT SUSHI 242.
 FT SUSHI 243.
 FT SUSHI 244.
 FT SUSHI 245.
 FT SUSHI 246.
 FT SUSHI 247.
 FT SUSHI 248.
 FT SUSHI 249.
 FT SUSHI 250.
 FT SUSHI 251.
 FT SUSHI 252.
 FT SUSHI 253.
 FT SUSHI 254.
 FT SUSHI 255.
 FT SUSHI 256.
 FT SUSHI 257.
 FT SUSHI 258.
 FT SUSHI 259.
 FT SUSHI 260.
 FT SUSHI 261.
 FT SUSHI 262.
 FT SUSHI 263.
 FT SUSHI 264.
 FT SUSHI 265.
 FT SUSHI 266.
 FT SUSHI 267.
 FT SUSHI 268.
 FT SUSHI 269.
 FT SUSHI 270.
 FT SUSHI 271.
 FT SUSHI 272.
 FT SUSHI 273.
 FT SUSHI 274.
 FT SUSHI 275.
 FT SUSHI 276.
 FT SUSHI 277.
 FT SUSHI 278.
 FT SUSHI 279.
 FT SUSHI 280.
 FT SUSHI 281.
 FT SUSHI 282.
 FT SUSHI 283.
 FT SUSHI 284.
 FT SUSHI 285.
 FT SUSHI 286.
 FT SUSHI 287.
 FT SUSHI 288.
 FT SUSHI 289.
 FT SUSHI 290.
 FT SUSHI 291.
 FT SUSHI 292.
 FT SUSHI 293.
 FT SUSHI 294.
 FT SUSHI 295.
 FT SUSHI 296.
 FT SUSHI 297.
 FT SUSHI 298.
 FT SUSHI 299.
 FT SUSHI 300.
 FT SUSHI 301.
 FT SUSHI 302.
 FT SUSHI 303.
 FT SUSHI 304.
 FT SUSHI 305.
 FT SUSHI 306.
 FT SUSHI 307.
 FT SUSHI 308.
 FT SUSHI 309.
 FT SUSHI 310.
 FT SUSHI 311.
 FT SUSHI 312.
 FT SUSHI 313.
 FT SUSHI 314.
 FT SUSHI 315.
 FT SUSHI 316.
 FT SUSHI 317.
 FT SUSHI 318.
 FT SUSHI 319.
 FT SUSHI 320.
 FT SUSHI 321.
 FT SUSHI 322.
 FT SUSHI 323.
 FT SUSHI 324.
 FT SUSHI 325.
 FT SUSHI 326.
 FT SUSHI 327.
 FT SUSHI 328.
 FT SUSHI 329.
 FT SUSHI 330.
 FT SUSHI 331.
 FT SUSHI 332.
 FT SUSHI 333.
 FT SUSHI 334.
 FT SUSHI 335.
 FT SUSHI 336.
 FT SUSHI 337.
 FT SUSHI 338.
 FT SUSHI 339.
 FT SUSHI 340.
 FT SUSHI 341.
 FT SUSHI 342.
 FT SUSHI 343.
 FT SUSHI 344.
 FT SUSHI 345.
 FT SUSHI 346.
 FT SUSHI 347.
 FT SUSHI 348.
 FT SUSHI 349.
 FT SUSHI 350.
 FT SUSHI 351.
 FT SUSHI 352.
 FT SUSHI 353.
 FT SUSHI 354.
 FT SUSHI 355.
 FT SUSHI 356.
 FT SUSHI 357.
 FT SUSHI 358.
 FT SUSHI 359.
 FT SUSHI 360.
 FT SUSHI 361.
 FT SUSHI 362.
 FT SUSHI 363.
 FT SUSHI 364.
 FT SUSHI 365.
 FT SUSHI 366.
 FT SUSHI 367.
 FT SUSHI 368.
 FT SUSHI 369.
 FT SUSHI 370.
 FT SUSHI 371.
 FT SUSHI 372.
 FT SUSHI 373.
 FT SUSHI 374.
 FT SUSHI 375.
 FT SUSHI 376.
 FT SUSHI 377.
 FT SUSHI 378.
 FT SUSHI 379.
 FT SUSHI 380.
 FT SUSHI 381.
 FT SUSHI 382.
 FT SUSHI 383.
 FT SUSHI 384.
 FT SUSHI 385.
 FT SUSHI 386.
 FT SUSHI 387.
 FT SUSHI 388.
 FT SUSHI 389.
 FT SUSHI 390.
 FT SUSHI 391.
 FT SUSHI 392.
 FT SUSHI 393.
 FT SUSHI 394.
 FT SUSHI 395.
 FT SUSHI 396.
 FT SUSHI 397.
 FT SUSHI 398.
 FT SUSHI 399.
 FT SUSHI 400.
 FT SUSHI 401.
 FT SUSHI 402.
 FT SUSHI 403.
 FT SUSHI 404.
 FT SUSHI 405.
 FT SUSHI 406.
 FT SUSHI 407.
 FT SUSHI 408.
 FT SUSHI 409.
 FT SUSHI 410.
 FT SUSHI 411.
 FT SUSHI 412.
 FT SUSHI 413.
 FT SUSHI 414.
 FT SUSHI 415.
 FT SUSHI 416.
 FT SUSHI 417.
 FT SUSHI 418.
 FT SUSHI 419.
 FT SUSHI 420.
 FT SUSHI 421.
 FT SUSHI 422.
 FT SUSHI 423.
 FT SUSHI 424.
 FT SUSHI 425.
 FT SUSHI 426.
 FT SUSHI 427.
 FT SUSHI 428.
 FT SUSHI 429.
 FT SUSHI 430.
 FT SUSHI 431.
 FT SUSHI 432.
 FT SUSHI 433.
 FT SUSHI 434.
 FT SUSHI 435.
 FT SUSHI 436.
 FT SUSHI 437.
 FT SUSHI 438.
 FT SUSHI 439.
 FT SUSHI 440.
 FT SUSHI 441.
 FT SUSHI 442.
 FT SUSHI 443.
 FT SUSHI 444.
 FT SUSHI 445.
 FT SUSHI 446.
 FT SUSHI 447.
 FT SUSHI 448.
 FT SUSHI 449.
 FT SUSHI 450.
 FT SUSHI 451.
 FT SUSHI 452.
 FT SUSHI 453.
 FT SUSHI 454.
 FT SUSHI 455.
 FT SUSHI 456.
 FT SUSHI 457.
 FT SUSHI 458.
 FT SUSHI 459.
 FT SUSHI 460.
 FT SUSHI 461.
 FT SUSHI 462.
 FT SUSHI 463.
 FT SUSHI 464.
 FT SUSHI 465.
 FT SUSHI 466.
 FT SUSHI 467.
 FT SUSHI 468.
 FT SUSHI 469.
 FT SUSHI 470.
 FT SUSHI 471.
 FT SUSHI 472.
 FT SUSHI 473.
 FT SUSHI 474.
 FT SUSHI 475.
 FT SUSHI 476.
 FT SUSHI 477.
 FT SUSHI 478.
 FT SUSHI 479.
 FT SUSHI 480.
 FT SUSHI 481.
 FT SUSHI 482.
 FT SUSHI 483.
 FT SUSHI 484.
 FT SUSHI 485.
 FT SUSHI 486.
 FT SUSHI 487.
 FT SUSHI 488.
 FT SUSHI 489.
 FT SUSHI 490.
 FT SUSHI 491.
 FT SUSHI 492.
 FT SUSHI 493.
 FT SUSHI 494.
 FT SUSHI 495.
 FT SUSHI 496.
 FT SUSHI 497.
 FT SUSHI 498.
 FT SUSHI 499.
 FT SUSHI 500.
 FT SUSHI 501.
 FT SUSHI 502.
 FT SUSHI 503.
 FT SUSHI 504.
 FT SUSHI 505.
 FT SUSHI 506.
 FT SUSHI 507.
 FT SUSHI 508.
 FT SUSHI 509.
 FT SUSHI 510.
 FT SUSHI 511.
 FT SUSHI 512.
 FT SUSHI 513.
 FT SUSHI 514.
 FT SUSHI 515.
 FT SUSHI 516.
 FT SUSHI 517.
 FT SUSHI 518.
 FT SUSHI 519.
 FT SUSHI 520.
 FT SUSHI 521.
 FT SUSHI 522.
 FT SUSHI 523.
 FT SUSHI 524.
 FT SUSHI 525.
 FT SUSHI 526.
 FT SUSHI 527.
 FT SUSHI 528.
 FT SUSHI 529.
 FT SUSHI 530.
 FT SUSHI 531.
 FT SUSHI 532.
 FT SUSHI 533.
 FT SUSHI 534.
 FT SUSHI 535.
 FT SUSHI 536.
 FT SUSHI 537.
 FT SUSHI 538.
 FT SUSHI 539.
 FT SUSHI 540.
 FT SUSHI 541.
 FT SUSHI 542.
 FT SUSHI 543.
 FT SUSHI 544.
 FT SUSHI 545.
 FT SUSHI 546.
 FT SUSHI 547.
 FT SUSHI 548.
 FT SUSHI 549.
 FT SUSHI 550.
 FT SUSHI 551.
 FT SUSHI 552.
 FT SUSHI 553.
 FT SUSHI 554.
 FT SUSHI 555.
 FT SUSHI 556.
 FT SUSHI 557.
 FT SUSHI 558.
 FT SUSHI 559.
 FT SUSHI 560.
 FT SUSHI 561.
 FT SUSHI 562.
 FT SUSHI 563.
 FT SUSHI 564.
 FT SUSHI 565.
 FT SUSHI 566.
 FT SUSHI 567.
 FT SUSHI 568.
 FT SUSHI 569.
 FT SUSHI 570.
 FT SUSHI 571.
 FT SUSHI 572.
 FT SUSHI 573.
 FT SUSHI 574.
 FT SUSHI 575.
 FT SUSHI 576.
 FT SUSHI 577.
 FT SUSHI 578.
 FT SUSHI 579.
 FT SUSHI 580.
 FT SUSHI 581.
 FT SUSHI 582.
 FT SUSHI 583.
 FT SUSHI 584.
 FT SUSHI 585.
 FT SUSHI 586.
 FT SUSHI 587.
 FT SUSHI 588.
 FT SUSHI 589.
 FT SUSHI 590.
 FT SUSHI 591.
 FT SUSHI 592.
 FT SUSHI 593.
 FT SUSHI 594.
 FT SUSHI 595.
 FT SUSHI 596.
 FT SUSHI 597.
 FT SUSHI 598.
 FT SUSHI 599.
 FT SUSHI 600.
 FT SUSHI 601.
 FT SUSHI 602.
 FT SUSHI 603.
 FT SUSHI 604.
 FT SUSHI 605.
 FT SUSHI 606.
 FT SUSHI 607.
 FT SUSHI 608.
 FT SUSHI 609.
 FT SUSHI 610.
 FT SUSHI 611.
 FT SUSHI 612.
 FT SUSHI 613.
 FT SUSHI 614.
 FT SUSHI 615.
 FT SUSHI 616.
 FT SUSHI 617.
 FT SUSHI 618.
 FT SUSHI 619.
 FT SUSHI 620.
 FT SUSHI 621.
 FT SUSHI 622.
 FT SUSHI 623.
 FT SUSHI 624.
 FT SUSHI 625.
 FT SUSHI 626.
 FT SUSHI 627.
 FT SUSHI 628.
 FT SUSHI 629.
 FT SUSHI 630.
 FT SUSHI 631.
 FT SUSHI 632.
 FT SUSHI 633.
 FT SUSHI 634.
 FT SUSHI 635.
 FT SUSHI 636.
 FT SUSHI 637.
 FT SUSHI 638.
 FT SUSHI 639.
 FT SUSHI 640.
 FT SUSHI 641.
 FT SUSHI 642.
 FT SUSHI 643.
 FT SUSHI 644.
 FT SUSHI 645.
 FT SUSHI 646.
 FT SUSHI 647.
 FT SUSHI 648.
 FT SUSHI 649.
 FT SUSHI 650.
 FT SUSHI 651.
 FT SUSHI 652.
 FT SUSHI 653.
 FT SUSHI 654.
 FT SUSHI 655.
 FT SUSHI 656.
 FT SUSHI 657.
 FT SUSHI 658.
 FT SUSHI 659.
 FT SUSHI 660.
 FT SUSHI 661.
 FT SUSHI 662.
 FT SUSHI 663.
 FT SUSHI 664.
 FT SUSHI 665.
 FT SUSHI 666.
 FT SUSHI 667.
 FT SUSHI 668.
 FT SUSHI 669.
 FT SUSHI 670.
 FT SUSHI 671.
 FT SUSHI 672.
 FT SUSHI 673.
 FT SUSHI 674.
 FT SUSHI 675.
 FT SUSHI 676.
 FT SUSHI 677.
 FT SUSHI 678.
 FT SUSHI 679.
 FT SUSHI 680.
 FT SUSHI 681.
 FT SUSHI 682.
 FT SUSHI 683.
 FT SUSHI 684.
 FT SUSHI 685.
 FT SUSHI 686.
 FT SUSHI 687.
 FT SUSHI 688.
 FT SUSHI 689.
 FT SUSHI 690.
 FT SUSHI 691.
 FT SUSHI 692.
 FT SUSHI 693.
 FT SUSHI 694.
 FT SUSHI 695.
 FT SUSHI 696.
 FT SUSHI 697.
 FT SUSHI 698.
 FT SUSHI 699.
 FT SUSHI 700.
 FT SUSHI 701.
 FT SUSHI 702.
 FT SUSHI 703.
 FT SUSHI 704.
 FT SUSHI 705.
 FT SUSHI 706.
 FT SUSHI 707.
 FT SUSHI 708.
 FT SUSHI 709.
 FT SUSHI 710.
 FT SUSHI 711.
 FT SUSHI 712.
 FT SUSHI 713.
 FT SUSHI 714.
 FT SUSHI 715.
 FT SUSHI 716.
 FT SUSHI 717.
 FT SUSHI 718.
 FT SUSHI 719.
 FT SUSHI 720.
 FT SUSHI 721.
 FT SUSHI 722.
 FT SUSHI 723.
 FT SUSHI 724.
 FT SUSHI 725.
 FT SUSHI 726.
 FT SUSHI 727.
 FT SUSHI 728.
 FT SUSHI 729.
 FT SUSHI 730.
 FT SUSHI 731.
 FT SUSHI 732.
 FT SUSHI 733.
 FT SUSHI 734.
 FT SUSHI 735.
 FT SUSHI 736.
 FT SUSHI 737.
 FT SUSHI 738.
 FT SUSHI 739.
 FT SUSHI 740.
 FT SUSHI 741.
 FT SUSHI 742.
 FT SUSHI 743.
 FT SUSHI 744.
 FT SUSHI 745.
 FT SUSHI 746.
 FT SUSHI 747.
 FT SUSHI 748.
 FT SUSHI 749.
 FT SUSHI 750.
 FT SUSHI 751.
 FT SUSHI 752.
 FT SUSHI 753.
 FT SUSHI 754.
 FT SUSHI 755.
 FT SUSHI 756.
 FT SUSHI 757.
 FT SUSHI 758.
 FT SUSHI 759.
 FT SUSHI 760.
 FT SUSHI 761.
 FT SUSHI 762.
 FT SUSHI 763.
 FT SUSHI 764.
 FT SUSHI 765.
 FT SUSHI 766.
 FT SUSHI 767.
 FT SUSHI 768.
 FT SUSHI 769.
 FT SUSHI 770.
 FT SUSHI 771.
 FT SUSHI 772.
 FT SUSHI 773.
 FT SUSHI 774.
 FT SUSHI 775.
 FT SUSHI 776.
 FT SUSHI 777.
 FT SUSHI 778.
 FT SUSHI 779.
 FT SUSHI 780.
 FT SUSHI 781.
 FT SUSHI 782.
 FT SUSHI 783.
 FT SUSHI 784.
 FT SUSHI 785.
 FT SUSHI 786.
 FT SUSHI 787.
 FT SUSHI 788.
 FT SUSHI 789.
 FT SUSHI 790.
 FT SUSHI 791.
 FT SUSHI 792.
 FT SUSHI 793.
 FT SUSHI 794.
 FT SUSHI 795.
 FT SUSHI 796.
 FT SUSHI 797.
 FT SUSHI 798.
 FT SUSHI 799.
 FT SUSHI 800.
 FT SUSHI 801.
 FT SUSHI 802.
 FT SUSHI 803.
 FT SUSHI 804.
 FT SUSHI 805.
 FT SUSHI 806.
 FT SUSHI 807.
 FT SUSHI 808.
 FT SUSHI 809.
 FT SUSHI 810.
 FT SUSHI 811.
 FT SUSHI 812.
 FT SUSHI 813.
 FT SUSHI 814.
 FT SUSHI 815.
 FT SUSHI 816.
 FT SUSHI 817.
 FT SUSHI 818.
 FT SUSHI 819.
 FT SUSHI 820.
 FT SUSHI 821.
 FT SUSHI 822.
 FT SUSHI 823.
 FT SUSHI 824.
 FT SUSHI 825.
 FT SUSHI 826.
 FT SUSHI 827.
 FT SUSHI 828.
 FT SUSHI 829.
 FT SUSHI 830.

FT CARBOHYD 145 145 N LINKED (GLCNAC...) (POTENTIAL)
 FT CARBOHYD 205 205 N LINKED (GLCNAC...) (POTENTIAL)
 EQ SEQUENCE 297 AA; 33197 MW; 911209DA6C119D59 CRC64;

Query Match 17.8%; Score 134; DB 1; Length 297;
 Best Local Similarity 26.6%; Pred. No. 2.4e-06;
 Matches 37; Conservative 19; Mismatches 65; Indels 18; Gaps 6;

QY 2 SCDEPPFVKNAKKPYSLP... VGLVLRVTCSPVRLGKKAIFGLSENOVHATW 56
 DB 92 HCPPIPPPKFAALKEYTSCVNSFFQDIIVFKCLPHFAMFGNDIVICTA-----HCNW 146
 QY 57 DKAPPICSVNKTISCSDPIVPGGPMNKGSKAPPHGNSVTFICKANFTMKSKTIVWQA 116
 DB 147 TQLPCKRE... VKCPDPSRDNCFENVYDAKPVLSYKDKAVPCCHTYKLDGPERVECT 201
 QY 116 ANEMGCTALPVCES 134
 DB 292 KTGHW SALLFSKASCKL 218

RESULT 28
 C400 RAT STANDARD; PRT; 258 AA.
 AC C64515;
 DT 01 NOV 1997 (rel. 45, last sequence update)
 DT 15 JUL 1998 (rel. 46, last annotation update)
 DE C4B binding protein beta chain precursor.
 GN C400B.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 OX NCBI_TaxID=10116;
 RN 111
 RP SEQUENCE FROM N.A.
 RC STRAIN: SPRAGUE DAWLEY; TISSUE: Liver;
 RX MEDLINE-97166082; PubMed-9011975;
 RA Hillarp A., Wiklund H., Thern A., Dahlback B.;
 RT "Molecular cloning of rat C4B binding protein alpha- and beta chains;
 RT structural and functional relationships among human, bovine, rabbit,
 RT mouse, and rat proteins.";
 RL J. Immunol. 158:1315-1323(1997).
 CC 1- FUNCTION: C4BP CONTROLS THE CLASSICAL PATHWAY OF COMPLEMENT
 CC ACTIVATION. IT BINDS AS A COFACTOR TO C3b/C4b INACTIVATOR
 CC (CHINA). WHICH THEN HYDROLYZES THE COMPLEMENT FRAGMENT C4b. IT
 CC ALSO ACCELERATES THE DEGRADATION OF THE C4bC2a COMPLEX (C3
 CC CONVERTASE) BY DISSOCIATING THE COMPLEMENT FRAGMENT C2a. IT
 CC INTERACTS ALSO WITH ANTICOGULANT PROTEIN S AND WITH SERUM AMYLOID
 CC P COMPONENT.
 CC 1- SUBUNIT: DISULFIDE-LINKED COMPLEX OF C4BP ALPHA AND BETA CHAINS.
 CC 1- SIMILARITY: TO C4BP ALPHA CHAIN AND TO PIG APOLIPOPROTEIN R.
 CC 1- SIMILARITY: CONTAINS 3 SUSHI (SCR) DOMAINS.

This SWISS PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (see <http://www.isb-sib.ch/announcements/>
 or send an email to license@isb-sib.ch).

EMBL: Z50052; CAA90192.1;
 EMBL: P10948; IVD0.
 EMBL: InterPro: IPR000436; Sushi_SCR_C4BP.
 EMBL: Pfam: PF00084; Sushi_1.
 EMBL: SMART: SM00042; C4BP_1.
 KW Complement pathway; Plasma; glycoprotein; Repeat; Sushi; Signal.
 FT SIGNAL 1 15 BY SIMILARITY.
 FT CHAIN 16 258 C4B BINDING PROTEIN BETA CHAIN.
 FT DOMAIN 19 74 SUSHI 1 (ATYPICAL; LACK A CYS).
 FT DOMAIN 77 132 SUSHI 2.
 FT DOMAIN 105 189 SUSHI 3.

FT DISULFID 46 73 BY SIMILARITY.
 FT DISULFID 78 118 BY SIMILARITY.
 FT DISULFID 136 176 BY SIMILARITY.
 FT DISULFID 162 188 BY SIMILARITY.
 FT DISULFID 217 217 INTERHA 4 (WITH ALPHA CHAIN)
 FT CARBOHYD 27 27 (POTENTIAL).
 FT CARBOHYD 67 67 N-LINKED (GLCNAC...) (POTENTIAL).
 FT CARBOHYD 89 89 N-LINKED (GLCNAC...) (POTENTIAL).
 FT CARBOHYD 95 95 N-LINKED (GLCNAC...) (POTENTIAL).
 FT CARBOHYD 114 114 N-LINKED (GLCNAC...) (POTENTIAL).
 FT CARBOHYD 218 218 N-LINKED (GLCNAC...) (POTENTIAL).
 SQ SEQUENCE 258 AA; 28641 MW; 6363999J8A450K1C CRC64;

Query Match 17.8%; Score 133 5; DB 1; Length 258;
 Best Local Similarity 26.18; Pred. No. 2.3e-06;
 Matches 35; Conservative 17; Mismatches 57; Indels 25; Gaps 5;

QY 3 SCDEPPFVKNAKKPYSLPVGIVLR-----VFGKFTPEQIIIG 113
 DB 19 SCSEPPPVNNS-----VFGKFTPEQIIIG 113
 QY 57 DKAPPICSVNKTISCSDPIVPGGPMNKGSKAPPHGNSVTFICKANFTMKSKTIVWQA 116
 DB 67 NSTPLPCLLGH---CPDPVLENGKIN--SSDPNISKRMFECNFGYILKSNWSQLE 120
 QY 117 NEMWGCTALPVCES 130
 DB 121 DHTWAP-LPICKS 133

RESULT 29
 LEM2 MOUSE STANDARD; PRT; 6.2 AA
 AC QC0690;
 DT 01-APR-1994 (rel. 25, created)
 DT 01-APR-1994 (rel. 25, last sequence update)
 DE E-selectin precursor (Endothelial leukocyte adhesion molecule 1)
 DE (E-AM-1) (leukocyte-endothelial cell adhesion molecule 2) (E-AM2)
 DE (CD62E).
 GN SELE OR ELAM-1.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN 111
 RP SEQUENCE FROM N.A.
 RX MEDLINE-92283265; PubMed-1375914;
 RA Becker-Andre M., van Halbeek H. Losberger C., Whelan J.,
 RA De Amater J.F.;
 RT "Murine endothelial leukocyte adhesion molecule 1 is a close
 RT structural and functional homologue of the human protein.";
 RL Eur. J. Biochem. 206:401-411(1992)
 RN 12
 RP SEQUENCE FROM N.A.
 RX MEDLINE-92240571; PubMed-1378846;
 RA Weier A., Isenmann S., Vestweber D.;
 RT "Cloning of the mouse endothelial selectin Expression of both E-
 RT and P-selectin is inducible by tumor necrosis factor alpha.";
 RL J. Biol. Chem. 267:15176-15183(1992).
 CC 1- FUNCTION: EXPRESSED ON CYTOKINE INDUCED ENDOTHELIAL CELLS AND
 CC MEDIATES THEIR BINDING TO LEUKOCYTES. THE LIGAND RECOGNIZED BY
 CC E-AM-1 IS SIALYL-LEWIS X (ALPHA(1-3)FUCOSYLATED DERIVATIVES OF
 CC POLYLACTOSAMINE THAT ARE FOUND AT THE NONREDUCING TERMINI OF
 CC GLYCOPROTEINS).
 CC 1- SUBCELLULAR LOCATION: Type I membrane protein.
 CC 1- SIMILARITY: TO OTHER SELECTINS/ECAMS.
 CC 1- SIMILARITY: CONTAINS 1 C-TYPE LECTIN FAMILY DOMAIN.
 CC 1- SIMILARITY: CONTAINS 1 EGF-LIKE DOMAIN.
 CC 1- SIMILARITY: CONTAINS 6 SUSHI (SCR) DOMAINS.

This SWISS-PROT entry is copyright. It is produced through a collaboration

group. ?
 BL Biochem. J. 213:201-209(1984).
 RN [8]
 RN VARIANTS C2D PRP-209 AND APO-464
 RP MEDLINE-96215049; PubMed-8621452;
 RA Wetzel R.A., Kallies J., Lakk M.I., Kiepiela P., Akama H.,
 RA Johnson C.A., Benson P., Collins H.P.;
 RT "Type II human complement C2 deficiency. Allele specific amino acid
 RT substitutions (Ser189 -> Phe; Gly444 -> Arg) cause impaired C2
 RT secretion.";
 RL J. Biol. Chem. 271:5824-5831(1996).
 RN [9]
 RN VARIANT C2D TYR-141.
 RP MEDLINE-96344005; PubMed-9670940;
 RA Zhu Z.H., Atkinson T.P., Volanakis J.E.;
 RA "A novel type II complement C2 deficiency allele in an African-
 RA American family.";
 RL J. Immunol. 161:578-584(1998).
 CC -1 FUNCTION: COMPONENT C2 WHICH IS PART OF THE CLASSICAL PATHWAY OF
 CC THE COMPLEMENT SYSTEM IS CLEAVED BY ACTIVATED FACTOR C1 INTO TWO
 CC FRAGMENTS: C2B AND C2A. C2A, A SERINE PROTEASE, THEN COMBINES WITH
 CC COMPLEMENT FACTOR 4B TO GENERATE THE C3 OR C5 CONVERTASE
 CC -1 CATALYTIC ACTIVITY: CLEAVES C3 IN THE ALPHA-CHAIN TO YIELD C3A AND
 CC C3B. CLEAVES C5 IN THE ALPHA-CHAIN TO YIELD C5A AND C5B. BOTH
 CC CLEAVAGES TAKE PLACE AT THE C-TERMINAL OF AN ARGININE RESIDUE.
 CC -1 DISEASE: DEFECTS IN C2 ARE THE CAUSE OF C2 DEFICIENCY (C2D). THIS
 CC IS AN AUTOSOMAL RECESSIVE DISEASE. DEFICIENT INDIVIDUALS HAVE AN
 CC INCREASED INCIDENCE OF SLE AND SLE-LIKE SYNDROMES,
 CC GLOMERULONEPHRITIS, VASCULITIS AND PYOGENIC INFECTIONS. TYPE I C2D
 CC IS CHARACTERIZED BY COMPLETE LOSS OF THE PROTEIN WHILE TYPE II C2D
 CC IS CHARACTERIZED BY A SELECTIVE BLOCK IN C2 SECRETION.
 CC -1 MISCELLANEOUS: C2 IS A MAJOR HISTOCOMPATIBILITY COMPLEX CLASS-III
 CC PROTEIN.
 CC -1 SIMILARITY: WITH COMPLEMENT FACTOR H.
 CC -1 SIMILARITY: CONTAINS 3 SUSHI (SCR) DOMAINS.
 CC -1 SIMILARITY: CONTAINS 1 VWFA DOMAIN.
 CC -1 SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1; ALSO KNOWN AS THE
 CC TRYPSIN FAMILY.
 CC
 CC This SWISS-Prot entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 DR EMBL: M15082; AAA59624.1; ?
 DR EMBL: L09708; AAB97607.1; ?
 DR EMBL: L09706; AAB97607.1; JOINED.
 DR EMBL: L09707; AAB97607.1; JOINED.
 DR EMBL: AF019411; AAB67975.1; ?
 DR EMBL: X04481; CAA28169.1; ?
 DR PIR: A25290; C2HU.
 DR PIR: A05289; A05289.
 DR HSSP: P00734; 1VR1.
 DR MEROPS: S01194; ?
 DR MIM: 217000;
 DR InterPro: IPK001114; Chymotrypsin.
 DR InterPro: IPK000436; Sushi_SCR_CCP.
 DR InterPro: IPK001254; Trypsin.
 DR InterPro: IPK002035; VWFA.
 DR Pfam: PF00084; sushi; 3.
 DR Pfam: PF00089; trypsin; 2.
 DR Pfam: PF00092; vwa; 1.
 DR PRINTS: PR00722; CHYMOTRYPSIN.
 DR SMART: SM00042; CCP; 4.
 DR SMART: SM00020; Tryp_Sp; 1.
 DR SMART: SM00427; VWA; 1.
 DR PROSITE: PS00240; TRYPSIN_DOM; 1.
 DR PROSITE: PS00146; TRYPSIN_HIS; 1.
 DR PROSITE: PS00145; TRYPSIN_SER; 1.
 DR PROSITE: PS00744; VWFA; 1

KW Complement pathway; Plasma; Glycoprotein; Hydrolase; Serine protease;
 KW Signal, Refeat, Sushi, Disease mutation; Polymorphism.
 FT SIGNAL 1 20
 FT CHAIN 21 752 COMPLEMENT C2.
 FT CHAIN 21 243 COMPLEMENT C2B FRAGMENT.
 FT CHAIN 214 752 COMPLEMENT C2A FRAGMENT.
 FT DOMAIN 23 85 SUSHI 1.
 FT DOMAIN 38 145 SUSHI 2.
 FT DOMAIN 150 205 SUSHI 3.
 FT DOMAIN 254 452 VWFA.
 FT DOMAIN 456 752 SERINE PROTEASE.
 FT ACT_SITE 507 507 CHARGE RELAY SYSTEM (BY SIMILARITY).
 FT ACT_SITE 561 561 CHARGE RELAY SYSTEM (BY SIMILARITY).
 FT ACT_SITE 679 679 CHARGE RELAY SYSTEM (BY SIMILARITY).
 FT DISULFID 24 64 BY SIMILARITY.
 FT DISULFID 51 84 BY SIMILARITY.
 FT DISULFID 89 131 BY SIMILARITY.
 FT DISULFID 117 144 BY SIMILARITY.
 FT DISULFID 151 191 BY SIMILARITY.
 FT DISULFID 177 204 BY SIMILARITY.
 FT CARBOHYD 29 29 N-LINKED GLCNAC. (POTENTIAL).
 FT CARBOHYD 112 112 N-LINKED GLCNAC. (POTENTIAL).
 FT CARBOHYD 290 290 N-LINKED GLCNAC. (POTENTIAL).
 FT CARBOHYD 333 333 N-LINKED GLCNAC. (POTENTIAL).
 FT CARBOHYD 467 467 N-LINKED GLCNAC. (POTENTIAL).
 FT CARBOHYD 471 471 N-LINKED GLCNAC. (POTENTIAL).
 FT CARBOHYD 621 621 N-LINKED GLCNAC. (POTENTIAL).
 FT CARBOHYD 651 651 N-LINKED GLCNAC. (POTENTIAL).
 FT VARIANT 131 131 C -> V; N C2D, TYPE II).
 FT /FTID-VAR:008544.
 FT VARIANT 209 209 S -> F; N C2D, TYPE II).
 FT /FTID-VAR:008545.
 FT VARIANT 454 464 G -> R; N C2D, TYPE II).
 FT /FTID-VAR:008546.
 FT VARIANT 533 533 F -> L; N DHSRI:1042664.
 FT /FTID-VAR:011772.
 FT SEQUENCE 752 AA; 83267 MW; 5A96AL3700CF444 CRC64;
 Query Match 17.48; Score 140.5; DR 1; Length 752;
 Best local Similarity 27.08; Pred. No 1.5e 05;
 Matches 38; Conservative 19; Mismatches 61; Indels 24; Gaps 8;
 QY 2 ISCDPPEVKNRPYY-----SLPIVPGVVLAVTCSPSYRLIGEKAIFCISENQVHATW 56
 DB 87 VCAIAVSEFNG---IYPIKIGSYV--GCVASVECDKGFILKGSIVRQC---RKNQMW 147
 QY 57 DRAPPICHSVAKTISCSDPIVPGGFMNKGSKADPRHGDVSFTTCANFTMGSKIVWFOA 116
 DB 138 DSEIAVCD--NVAQSHENPPIISLAVATGPE---GHRKVPVYPPSSNIVLTSSSEPEQWS 143
 QY 117 NPMNGHTALPVCHS-----DPP 133
 DB 194 NGVMSCTE-PICRQPPSYDEP 213
 RESULT 11
 CFAR_MOUSE STANDARD; PRT; 161 AA
 ID CFAR_MOUSE
 AC P04186;
 DT 20-MAR-1987 (Rel. 04, Created)
 DT 01-AUG-1991 (Rel. 19, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Complement factor B precursor (pC3.4.4.47). (C3/C5 convertase).
 GN BF OR H2-HF.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Euteleia; Rodentia; Sciurotopathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID-10090;
 RN 11
 RP SEQUENCE FROM N.A.
 RX MEDLINE-91035130; PubMed-2229060;
 RA Ishikawa N., Nonaka M., Wetzel R.A.; (often H.R.);
 RT "Murine complement C2 and factor B genes and cDNA cloning reveals


```

[1]
SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RX MEDLINE-91177916; PubMed-2007602;
RA Mota T., Miyata T., Misumi Y., Tokunaga F., Nakamura T., Toh Y.,
RA Ikohata Y., Iwanaga S.;
RT "Limulus factor C: An endotoxin sensitive serine protease zymogen
RT with a mosaic structure of complement-like, epidermal growth
RT factor-like, and lectin-like domains.";
RE J. Biol. Chem. 266:6554-6561(1991).
CC 1 FUNCTION: THIS ENZYME IS CLOSELY ASSOCIATED WITH AN ENDOTOXIN-
CC SENSITIVE HEMOLYMPH COAGULATION SYSTEM WHICH MAY PLAY IMPORTANT
CC ROLES IN BOTH HEMOSTASIS AND HOST DEFENSE MECHANISMS. ITS ACTIVE
CC FORM CATALYZES THE ACTIVATION OF FACTOR B.
CC 2 CATALYTIC ACTIVITY: Selective cleavage of 103-Arg-1-Ser-104 and
CC 124-Ile-111e 125 bonds in Limulus clotting factor B to form
CC activated factor B. Cleavage of Pro-Arg-1 Xaa bonds in synthetic
CC substrates.
CC 3 ENZYME REGULATION: ACTIVATED BY GRAM-NEGATIVE BACTERIAL
CC LIPOLYSACCHARIDES AND CHYMOTRYPSIN.
CC 4 SUBUNIT: HETERODIMER OF A LIGHT CHAIN AND A HEAVY CHAIN LINKED
CC BY A DISULFIDE BOND.
CC 5 SIMILARITY: CONTAINS 5 SUSHI (SCR) DOMAINS.
CC 6 SIMILARITY: CONTAINS 1 C-TYPE LECTIN FAMILY DOMAIN.
CC 7 SIMILARITY: CONTAINS 1 EGF-LIKE DOMAIN.
CC 8 SIMILARITY: CONTAINS 1 EGF-LIKE DOMAIN.
CC 9 SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1; ALSO KNOWN AS THE
CC TRYPSIN FAMILY.
CC
CC This SWISS PROT entry is copyright It is produced through a collaboration
CC between the SWISS Institute of Bioinformatics and the EMBL-outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
DR PM01: D90271; BAA1415.1;
DR PM02: D90272; BAA1416.1;
DR F01: A08748; A08748;
DR F02: B08748; B08748;
DR B03P: P00744; I0V0;
DR M000S: S01229;
DR InterPro: IP0001414; chymotrypsin;
DR InterPro: IP0000561; EGF-like;
DR InterPro: IP0004043; LECT;
DR InterPro: IP0000436; SUSHI_SPK_ECP;
DR InterPro: IP0001254; Trypsin;
DR InterPro: IP0001404; lectin C;
DR Pfam: PF00059; lectin c; 1;
DR Pfam: PF00084; sush1; 5;
DR Pfam: PF00089; trypsin; 1;
DR PRINTS: PR00722; CHYMOTRYPSIN;
DR SMART: SM00042; CCP; 5;
DR SMART: SM00044; CLEF1; 1;
DR SMART: SM00181; EGF; 1;
DR SMART: SM00020; TRYP-SPE; 1;
DR PROSITE: PS00615; C-TYPE-LECTIN-1; FALSE_NEG;
DR PROSITE: PS00041; C-TYPE-LECTIN-2; 1;
DR PROSITE: PS00022; EGF-1; 1;
DR PROSITE: PS01186; EGF-2; FALSE_NEG;
DR PROSITE: PS00420; EGF-3; 1;
DR PROSITE: PS00240; TRYPSIN_DOM; 1;
DR PROSITE: PS00144; TRYPSIN_BIS; 1;
DR PROSITE: PS00145; TRYPSIN_SER; 1;
KW Hydrolase; Serine protease; Signal; Alternative splicing; Lectin;
KW hemolymph clotting; Glycoprotein; Cell adhesion; EGF-like domain;
KW SUSHI; Repeat; 1 25
FT CHAIN 26 1019 LIMULUS CLOTTING FACTOR C;
FT CHAIN 26 690 LIMULUS CLOTTING FACTOR C; HEAVY CHAIN.
FT CHAIN 691 1019 LIMULUS CLOTTING FACTOR C; LIGHT CHAIN.
FT CHAIN 691 762 LIMULUS CLOTTING FACTOR C; A CHAIN.
FT CHAIN 763 1019 LIMULUS CLOTTING FACTOR C; B CHAIN.

```

```

FT DOMAIN 102 137 EGF-LIKE.
FT DOMAIN 142 195 SUSHI .
FT DOMAIN 200 254 SUSHI .
FT DOMAIN 260 321 SUSHI .
FT DOMAIN 325 421 LECT.
FT DOMAIN 436 568 C-TYPE-LECTIN.
FT DOMAIN 576 634 SUSHI .
FT DOMAIN 685 748 SUSHI .
FT DOMAIN 763 1019 SERINE PR TEASE.
FT ACT_SITE 809 809 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT ACT_SITE 865 865 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT BINDING 966 966 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT DOMAIN 643 689 SUBSTRATE (BY SIMILARITY).
FT DISULFID 106 118 PRO-RULE.
FT DISULFID 125 125 BY SIMILARITY.
FT DISULFID 127 136 BY SIMILARITY.
FT DISULFID 436 447 BY SIMILARITY.
FT DISULFID 464 564 BY SIMILARITY.
FT DISULFID 548 556 BY SIMILARITY.
FT DISULFID 794 810 BY SIMILARITY.
FT DISULFID 942 951 BY SIMILARITY.
FT DISULFID 962 996 BY SIMILARITY.
FT CARBOHYD 523 523 N-LINKED GLCNAC . . . (POTENTIAL).
FT CARBOHYD 534 534 N-LINKED GLCNAC . . . (POTENTIAL).
FT CARBOHYD 624 624 N-LINKED GLCNAC . . . (POTENTIAL).
FT CARBOHYD 740 740 N-LINKED GLCNAC . . . (POTENTIAL).
FT CARBOHYD 767 767 N-LINKED GLCNAC . . . (POTENTIAL).
FT CARBOHYD 912 912 N-LINKED GLCNAC . . . (POTENTIAL).
FT VAPSPIC 492 498 LITWIS - TUNICAT (IN SHORT ISOFORM).
FT VAPSPIC 499 1019 MISSING (IN SHORT ISOFORM).
SQ SEQUENCE 1019 AA; 112346 MW; 5B2F4CE71289B CRC64;
Query Match 17.4%; Score 140.5; Db 1; Length 1019;
Best Local Similarity 26.6%; Pred. No. 2.1e 05;
Matches 42; Conservative 22; Mismatches 57; Idents 47; Gaps 9;
QY 2 ISCHP-----PP-----EVKNARKYY- SLFTVGVIVRYTCTSYRELDEKAL 44
||| | | | | | | | | | | | | | | | | |
DB 180 ISCLPGRWS-PPPKTPETAKVSPPHBYVM-SNMFGALIPESDPSYVLTQETL 49
QY 45 FCISLNOVHAIWKAPPCICESVNKILSCS DP----VQGFPMKNGSKALPRHDSV 96
| | | | | | | | | | | | | | | | | | |
DB 240 TC-----QNGUWSGQIPQCK---KLIVFCHLDP/NHAFHGVKICVQKYQFP--Q 190
QY 97 TETCKANITMGSKITVWCQANEMWG---PTALPVCESD 141
||| | | | | | | | | | | | | | | | | |
DB 291 TVTCSNYFLMGFNILKCNPDGWSGSGQPS-VLAADRE 128
RESULT 33
DAF_PONY
ID DAF_PONY STANDARD; PPT; (10 AA.
AC P49457;
DT 01-FEB-1994 (rel. 33, Created)
DT 01-FEB-1994 (rel. 33, Last sequence update)
DE Complement decay-accelerating factor (CD55) (Fragment).
GN DAF OR CD55.
OS Homo pygmaeus (Orangutan).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Insecta; Primates; Catarrhi; Hominoidea; Hominidae;
OX NCBI_TaxID:96060;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE-9110622; PubMed-7506743;
RA McKells M W., Alvarez J. J., Lablitz O. A., Atkinson J. P.;
RT "Characterization of DAF-2, a high molecular weight form of decay
RT accelerating factor (DAF; CD55), as a covalently cross linked dimer
RT of DAF-1."
RL J. Immunol. 152:676-685(1994).
CC 1- FUNCTION: THIS PROTEIN RECOGNIZES 34H AAV C9B FRAGMENTS THAT
CC CONDENSE WITH CELL-SURFACE HYDROX. OR AMINO GROUPS WHEN NAG-2ENI

```


01 MAR 2002 (rel. 41, Last annotation update)
 DE Limulus clotting factor C precursor (P03421.04) (P03)
 OS Carinioscorpius rotundicauda (Southeast Asian horseshoe crab)
 CC Eukaryota; Metazoa; Arthropoda; Chelicerata; Merostomata; Xiphosura;
 CC Limulidae; Carinioscorpius.
 XX NBI TaxID=6048;
 [1]
 RP SEQUENCE FROM N.A.
 RX LSSUE=Blood;
 RX MEDLINE=952648506; PubMed=7594401;
 RX J Biol. Navas M.A. 111, No 4;
 RI "Molecular cloning and sequence analysis of factor C cDNA from the
 RI Singapore horseshoe crab, Carinioscorpius rotundicauda.";
 RL Mol. Biol. Biotechnol. 4:90-103(1995).
 CC 1. FUNCTION: THIS ENZYME IS CLOSELY ASSOCIATED WITH AN ENDOTOXIN-
 CC SENSITIVE HEMOLYMPH COAGULATION SYSTEM WHICH MAY PLAY IMPORTANT
 CC ROLES IN BOTH HEMOLYSIS AND HOST DEFENSE MECHANISMS. ITS ACTIVE
 CC FORM CATALYZES THE ACTIVATION OF FACTOR B
 CC CATALYTIC ACTIVITY: Selective cleavage of 103-Arg-1-Ser-104 and
 CC 124-Ile-1-Ile-125 bonds in Limulus clotting factor B to form
 CC activated factor B. Cleavage of Pro-Arg-1-Xaa bonds in synthetic
 CC substrates.
 CC 2. ENZYME REGULATION: ACTIVATED BY GRAM NEGATIVE BACTERIAL
 CC LIPOLYSACCHARIDES AND CHYMOTRYPSIN (BY SIMILARITY).
 CC 3. SUBUNIT: HETERODIMER OF A LIGHT CHAIN AND A HEAVY CHAIN LINKED
 CC BY A DISULFIDE BOND (BY SIMILARITY).
 CC 4. SIMILARITY: CONTAINS 5 SUSHI (SCR) DOMAINS.
 CC 5. SIMILARITY: CONTAINS 1 C-TYPE LECTIN FAMILY DOMAIN.
 CC 6. SIMILARITY: CONTAINS 1 EGF-LIKE DOMAIN.
 CC 7. SIMILARITY: CONTAINS 1 LECTIN DOMAIN.
 CC 8. SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1; ALSO KNOWN AS THE
 CC TRYPSIN FAMILY.

THIS SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (see [http://www.isb-sib.ch/announcements/](http://www.isb-sib.ch/announcements)
 or send an email to license@isb-sib.ch).

EMBL: S7064; AAB34361.1;
 BSSP: P0074; J000.
 MERRIPS: S01219;
 InterPro: IP0001314; Chymotrypsin.
 InterPro: IP0000961; EGF-like
 InterPro: IP0004043; Lect.
 InterPro: IP000436; Sushi_SCR_CCP.
 InterPro: IP001254; Trypsin.
 InterPro: IP001304; Lectin_C.
 Pfam: PF00054; Lectin_C; 1.
 Pfam: PF00084; Sushi; 5.
 Pfam: PF00087; Trypsin; 1.
 PRINTS: PR00722; CHYMOTRYPSIN.
 SMART: SM00042; CCP; 5.
 SMART: SM00044; CLECT; 1.
 SMART: SM00181; EGF; 1.
 SMART: SM00020; TRYPSIN; 1.
 PROSITE: PS00615; C-TYPE_LECTIN_1; FALSE_NEG.
 PROSITE: PS00041; C-TYPE_LECTIN_2; 1.
 PROSITE: PS00022; EGF_1; 1.
 PROSITE: PS01186; EGF_2; FALSE_NEG.
 PROSITE: PS00820; LECT_1.
 PROSITE: PS02400; TRYPSIN_DOM; 1.
 PROSITE: PS00134; TRYPSIN_HIS; 1.
 PROSITE: PS00135; TRYPSIN_SER; 1.
 HYDROLASE: Serine protease; Signal; Lectin; Hemolymph clotting;
 KW glycoprotein; Cell adhesion; EGF-like domain; Sushi; Repeat.
 FT SIGNAL 1 25 BY SIMILARITY.
 FT CHAIN 26 1019 LIMULUS CLOTTING FACTOR C.
 FT CHAIN 26 690 LIMULUS CLOTTING FACTOR C, HEAVY CHAIN.
 FT CHAIN 691 1019 LIMULUS CLOTTING FACTOR C, LIGHT CHAIN.
 FT CHAIN 691 762 LIMULUS CLOTTING FACTOR C, A CHAIN.

FT CHAIN 763 1019 LIMULUS CLOTTING FACTOR C, B CHAIN.
 FT DOMAIN 102 137 EGF-LIKE.
 FT DOMAIN 142 195 SUSHI 1.
 FT DOMAIN 200 254 SUSHI 2.
 FT DOMAIN 260 321 SUSHI 3.
 FT DOMAIN 325 421 SUSHI 4.
 FT DOMAIN 436 568 C-TYPE LECTIN
 FT DOMAIN 576 634 SUSHI 5.
 FT DOMAIN 685 748 SUSHI 6.
 FT DOMAIN 763 1019 SERINE PROTEASE.
 FT ACT_SITE 809 809 CHARGE RELAY SYSTEM (BY SIMILARITY).
 FT ACT_SITE 865 865 CHARGE RELAY SYSTEM (BY SIMILARITY).
 FT ACT_SITE 966 966 CHARGE RELAY SYSTEM (BY SIMILARITY).
 FT BINDING 960 960 SUBSTRATE (BY SIMILARITY).
 FT DOMAIN 643 689 PRO-RCH.
 FT DISULFID 106 118 BY SIMILARITY.
 FT DISULFID 112 125 BY SIMILARITY.
 FT DISULFID 127 136 BY SIMILARITY.
 FT DISULFID 436 447 BY SIMILARITY.
 FT DISULFID 464 564 BY SIMILARITY.
 FT DISULFID 538 556 BY SIMILARITY.
 FT DISULFID 794 810 BY SIMILARITY.
 FT DISULFID 932 951 BY SIMILARITY.
 FT CARBOHYD 962 996 N-LINKED (GLNNA2;) (POTENTIAL).
 FT CARBOHYD 523 523 N-LINKED (GLNNA2;) (POTENTIAL).
 FT CARBOHYD 534 534 N-LINKED (GLNNA2;) (POTENTIAL).
 FT CARBOHYD 624 624 N-LINKED (GLNNA2;) (POTENTIAL).
 FT CARBOHYD 740 740 N-LINKED (GLNNA2;) (POTENTIAL).
 FT CARBOHYD 767 767 N-LINKED (GLNNA2;) (POTENTIAL).
 FT CARBOHYD 912 912 N-LINKED (GLNNA2;) (POTENTIAL).
 SQ SEQUENCE 1019 AA; 112429 MW; 5187.ED8B17B6C3 CRC64;
 Query Match 17.0%; Score 12.5; BB 1; Length 1019;
 Best Local Similarity 24.7%; Pred NC 4.1e 05;
 Matches 39; Conservative 28; M-Smashes 54; Indels 47; Gaps 8;
 QY 2 ISCDF-----PPEV-----KNARKAYSLPVPETVLRYTCSYKLGKAL 44
 DE BRQ ISLNGQWSNPPKCIIECAVSSPFGKVNALSCNIECATIRFSCHSPYVYLGKEL 219
 QY 45 FCISINOVHAYWDKAPICESVNKISCS--DPI---VPGCFMKKSKAFPRHDSV 96
 DE 240 TC-----QINQWNGIPOCKNI---VFQPDLDVNHAEHKVIGVEKYGFPP--QSTEV 240
 QY 97 IPTCFANFTMGSKTVWCQANMNC---ETALVYCESD 131
 DE 291 TVTCSGNVFLMGFDILKCNPDGWSNSQ--SQKQVADSE 428
 RESULT 38
 ID APHL_HOVIN STANDARD; PRT; 345 AA
 AC A17690; Q58052;
 DT 01-AUG-1990 (Rel. 15, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 15-OCT-2001 (Rel. 40, Last annotation update)
 DE Beta-2-glycoprotein I precursor (Apolipoprotein H) (Apo H) (H2GPI)
 DE (Beta(2)GPI).
 GN APDH.
 OS Bos taurus (Bovine).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae.
 OC Bovidae; Bovinae; Bos.
 OX NCBI_TaxID=5913;
 RN 1
 PP SEQUENCE FROM N.A.
 RC TISSUE=Liver.
 RA Gao B., Virmani M., Pomm E., Lazar W.-Sley E., Sakauchi K.,
 RA Appella E., Kuros G., Takacs L.;
 RA Submitted (1992-1992) to the EMBL/Genbank/DBJ databases
 RF SEQUENCE 38 4-345 FROM N.A., PARTIAL SEQUENCE. AND DISULFIDE BONDS.
 RC TISSUE=Liver;

RN REVIEW
 RX MEDLINE=94226679; PubMed 8172644;
 RA Ziptel P.F., Skerka C.;
 RT "complement factor H and related proteins: an expanding family of
 complement-regulatory proteins?";
 RL Immunol. Today 15:121-126(1994).
 CC -! FUNCTION: MIGHT BE INVOLVED IN COMPLEMENT REGULATION. CAN
 CC ASSOCIATE WITH LIPIDPROTEINS AND MAY PLAY A ROLE IN LIPID
 CC METABOLISM.
 CC -! SUBCELLULAR LOCATION: Extracellular.
 CC -! TISSUE SPECIFICITY: LIVER.
 CC -! PTM: N-GLYCOSYLATED. TWO FORMS ARE OBSERVED; ONE WITH A SINGLE
 CC SIDE CHAIN AND THE OTHER WITH TWO.
 CC -! SIMILARITY: CONTAINS 5 SUSHI (SCR) DOMAINS.
 CC -! SIMILARITY: STRONG, TO FACTOR H.
 CC
 CC This SWISS PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed, usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 DR EMBL: M65292; AAA35946.1;
 DR EMBL: M65293; AAA35947.1;
 DR EMBL: A1049741; CAB53064.1;
 DR EMBL: X56209; CAA39066.1;
 DR FIC: A40455; A40455;
 DR FIC: S14604; S14604;
 DR EESP: P10998; 1VDV.
 DR MIM: 144371;
 DR InterPro: IPR000436; Sushi_SCR_CCP.
 DR Pfam: PF00084; sushi; 5.
 DR SMART: SM00042; CCP; 5.
 KW Repeat: Glycoprotein; Sushi: Signal; Polymorphism.
 FT SIGNAL 1 18
 FT CHAIN 19 440
 FT DOMAIN 22 84
 FT DOMAIN 86 141
 FT DOMAIN 146 202
 FT DOMAIN 207 264
 FT DOMAIN 265 328
 FT CARBOHYD 126 126
 FT CARBOHYD 194 194
 FT VARIANT 157 157
 FT VARIANT 159 159
 FT VARIANT 175 175
 FT VARIANT 71 71
 FT CONFLICT 71 71
 FT SEQUENCE 440 AA; 37661 MW; 8DCDD4F92A85E035 CRC64;
 Query Match 16.7% Score 125.5; DB:1; Length 430;
 Best Local Similarity 26.5%; Pred. No. 1,9e-05;
 Matches 45; Conservative 25; Mismatches 59; Indels 13; Gaps 7;
 QY 1 ELSVPPPEVKNARKPYSLTIVP GTVLRYSVPSVRLIGERKAIQISENQVHATWDKA 59
 D 144 DTSVNNPTVQNAVIVSRQMSKYISGERVYQCRSPYEMFGDEEVMGLNGN-----WTE- 197
 QY 60 PPTCSVAKVLISCS DPTIVGDFPMNKISKAPRHCDSVFTICKANFTMKSKIVWCQANE 118
 D 198 PPQKC DSTGKGPPPPIDNDQITSEPLSVYAPASSVVEYQCNLYQLEGNKRITCR-NG 254
 QY 119 MWG PTALPG 128
 D 255 QWSHPPKTLHPC 266
 RESULT 40

LEM3 BOVIN
 ID LEM3 BOVIN STANDARD; PRT: 646 AA.
 AC P42201;
 DT 01-NOV-1995 (rel. 32, Created)
 DT 01-NOV-1995 (rel. 32, Last sequence update)
 DT 15-JUL-1999 (rel. 38, Last annotation update)
 DE P-selectin precursor (Granule membrane protein 140) (GMP-140) (PAISLEM)
 DE (CD62P) (Leukocyte-endothelial cell adhesion molecule 3) (LECAM3).
 GN S-IP.
 OS Bos taurus (Bovine).
 CC Eukaryota; Metazoa; Chordata; Crariata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Cetartiodactyla; Luminantia; Pecora; Boviodae;
 CC Bovidae; Bovinae; Bos.
 CC NCBI_TaxID=9913;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Capillary endothelium;
 RX MEDLINE=93249394; PubMed-7683458;
 RA Strubel N.A., Nguyen M., Kansas G.J., Tedder T.F., Fischell J.;
 RT "Isolation and characterization of a bovine cDNA encoding a
 RT functional homolog of human p-selectin";
 RL Biochem. Biophys. Res. Commun. 192:334-344(1993).
 CC -! FUNCTION: CA(2+)-DEPENDENT RECEPTOR FOR MYELOID CELLS THAT BINDS
 CC TO CARBOHYDRATES ON NEUTROPHILS AND MONOCYTES. MEDIATES THE
 CC INTERACTION OF ACTIVATED ENDOTHELIAL CELLS OR PLATELETS WITH
 CC LEUKOCYTES. THE LIGAND RECOGNIZED IS SIALYL-LEWIS X.
 CC -! SUBCELLULAR LOCATION: TYPE I membrane protein.
 CC -! TISSUE SPECIFICITY: STORED IN WE ALPHA-GRANULES OF PLATELETS
 CC AND WEIBEL PALADE BODIES OF ENDOTHELIAL CELLS. UPON CELL
 CC ACTIVATION BY AGONISTS, P-SELECTIN IS TRANSPORTED RAPIDLY TO
 CC THE CELL SURFACE.
 CC -! SIMILARITY TO OTHER SELECTINS/LECAMS.
 CC -! SIMILARITY: CONTAINS 1 C-TYPE LIX IN FAMILY DOMAIN.
 CC -! SIMILARITY: CONTAINS 1 EGF-LIKE DOMAIN.
 CC -! SIMILARITY: CONTAINS 6 SUSHI (SCR) DOMAINS; BOVINE P-LECTIN LACKS
 CC THE HUMAN SUSHI-3, -4 AND -7 EQUIVALENTS.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed, usage by and for commercial
 CC entities requires a license agreement (See [http://www.isb-sib.ch/](http://www.isb-sib.ch/announce/)
 CC or send an email to license@isb-sib.ch).
 CC
 DR EMBL: U12641; AAA30743.1;
 DR HSP: P16109; IPSH.
 DR InterPro: IPR000561; EGF-like.
 DR InterPro: IPR002396; Selectin.
 DR InterPro: IPR000436; Sushi_SCR_CCP.
 DR InterPro: IPR001304; lectin_C.
 DR Pfam: PF00054; EGF; 1.
 DR Pfam: PF00084; sushi; 6.
 DR Pfam: PF00084; sushi; 6.
 DR PRINTS: PF00043; SELECTIN.
 DR SMART: SM00042; CCP; 6.
 DR SMART: SM00043; EGF; 1.
 DR SMART: SM00141; EGF; 1.
 DR PROSITE: PS03022; EGF_1; 1.
 DR PROSITE: PS01186; EGF_2; 1.
 DR PROSITE: PS03615; C-TYPE LECTIN_1;
 DR PROSITE: PS03041; C-TYPE LECTIN_2;
 KW Cell adhesion; transmembrane; Glycoprotein; EGF-like domain; Lectin;
 FT SIGNAL 1 41
 FT CHAIN 42 646
 FT DOMAIN 42 587
 FT TRANSMEM 588 611
 FT DOMAIN 612 646
 FT DOMAIN 58 158
 FT DOMAIN 159 195
 FT DOMAIN 199 258
 FT DOMAIN 261 320


```

DR      250 KGTW--SFLPTCRESKL 266
RES01.4.2
LEM2_PIG
ID      LEM2_PIG          STANDARD:      PRT:      484 AA.
AC      P08110;
DT      01-FEB-1996 (Rel. 33, Created)
DI      01-FEB-1996 (Rel. 33, last sequence update)
DI      15-JUL-1998 (Rel. 36, last annotation update)
DE      E-selectin precursor (Endothelial leukocyte adhesion molecule 1)
DE      (ELAM-1) (Leukocyte-endothelial cell adhesion molecule 2) (LECAM2)
DE      (CD62E).
DN      SELE.
OS      Sus scrofa (Pig).
OC      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC      Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
OX      NCBI_TaxID=9923;
RN      [1]
RP      SEQUENCE FROM N.A.
RX      TISSUE-Aortic endothelium;
RX      MEDLINE=95071192; PubMed=7526854;
RA      Rollins S.A., Evans M.J., Johnson K.K., Elliot E.A., Squinto S.P.,
RA      Mattis L.A., Rother R.P.;
RT      *Molecular and functional analysis of porcine E-selectin reveals a
RT      potential role in xenograft rejection.*;
RL      Biochem. Biophys. Res. Commun. 204:763-771(1994).
RN      [2]
RP      SEQUENCE FROM N.A.
RX      TISSUE-Aortic endothelium;
RX      MEDLINE=94271236; PubMed=7516159;
RA      Isang V.T.M., Haskard D.O., Robinson M.K.;
RT      *Cloning and expression kinetics of porcine vascular cell adhesion
RT      molecule.*;
RL      Biochem. Biophys. Res. Commun. 201:805-805(1994).
CC      1- FUNCTION: EXPRESSED ON CYTOKINE INDUCED ENDOTHELIAL CELLS AND
CC      MEDIATES THEIR BINDING TO LEUKOCYTES. THE LIGAND RECOGNIZED BY
CC      ELAM-1 IS STALKY-DIWEIS X (ALPHA(1->3)FUCCOSYLATED DERIVATIVES OF
CC      POLYGLYCOSAMINE THAT ARE FOUND AT THE NONREDUCING TERMINI OF
CC      GLYCOPROTEINS).
CC      1- FUNCTION: PLAYS AN IMPORTANT ROLE IN ACUTE CELLULAR ALLOGRAFT
CC      REJECTION AND PROBABLY ALSO IN XENOGRAFT REJECTION.
CC      1- SUBCELLULAR LOCATION: Type 1 membrane protein.
CC      1- SIMILARITY: TO OTHER SELECTINS/LECAMS.
CC      1- SIMILARITY: CONTAINS 1 C-TYPE LECTIN FAMILY DOMAIN
CC      1- SIMILARITY: CONTAINS 4 SUSHI (SCR) DOMAINS.
CC      1- SIMILARITY: CONTAINS 1 EGF-LIKE DOMAIN.
CC      1- THE HUMAN SUSHI-1 AND -4 EQUIVALENTS.
CC
CC      THIS SWISS-PROT entry is copyrighted. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (See http://www.isb.sib.ch/announce/
CC      or send an email to license@isb.sib.ch).
CC
DR      EMBL: L39076; AAA61545.1;
DR      EMBL: 008450; AAA21541.1;
DR      HSSP: P16581; IESL.
DR      InterPro: IPR000561; EGF-like.
DR      InterPro: IPR002196; Selectin.
DR      InterPro: IPR000436; Sush1_SCR_CCP.
DR      InterPro: IPR001304; lectin_C.
DR      Pfam: PF00008; EGF; 1.
DR      Pfam: PF00059; lectin_c; 1.
DR      Pfam: PF00084; SUSHI; 4.
DR      PRINTS: PK00343; SELECTIN.
DR      SMART: SM00042; CCP; 4.
DR      SMART: SM00044; CLECT; 1.
DR      SMART: SM00181; EGF; 1.
DR      PROSITE: PS00022; EGF_1; 1.
DR      PROSITE: PS01186; EGF_2; 1.
DR      PROSITE: PS00615; C-TYPE LECTIN_1; 1.
DR      PROSITE: PS50041; C-TYPE LECTIN_2; 1.
KW      Cell adhesion, Transmembrane, Glycoprotein, EGF-like domain, Lectin.
KW      Selectin, Signal, Sush1, Repeat.
FT      SIGNAL          1..22
FT      CHAIN           23..484
FT      DOMAIN          23..429
FT      TRANSMEM      430..451
FT      DOMAIN          452..484
FT      DOMAIN          39..139
FT      DOMAIN          140..176
FT      DOMAIN          180..236
FT      DOMAIN          239..299
FT      DOMAIN          302..362
FT      DOMAIN          365..421
FT      DISULFID        41..139
FT      DISULFID        112..131
FT      DISULFID        144..155
FT      DISULFID        149..164
FT      DISULFID        166..175
FT      DISULFID        181..222
FT      DISULFID        208..235
FT      DISULFID        240..285
FT      DISULFID        271..298
FT      DISULFID        303..348
FT      DISULFID        334..361
FT      DISULFID        366..407
FT      DISULFID        393..420
FT      CARBOHYD        61..61
FT      CARBOHYD        65..65
FT      CARBOHYD        79..79
FT      CARBOHYD        160..160
FT      CARBOHYD        201..201
FT      CARBOHYD        254..254
FT      CARBOHYD        376..376
FT      CARBOHYD        400..400
FT      CONFLICT        253..253
FT      CONFLICT        313..313
FT      CONFLICT        321..321
FT      CONFLICT        327..327
FT      CONFLICT        363..363
FT      CONFLICT        384..384
FT      CONFLICT        461..461
SQ      SEQUENCE 484 AA; 52567 MW; AFF24F25C1FD013 CRG64;
Query Match      16.48; Score 123.5; DB 1; Length 484;
Best Local Similarity 23.84; Pred. No. 4.6e-05;
Matches 34; conservative 18; Mismatches 62; Inlets 29; Gaps 5;
QY      2 ISCDP DP-VKNARKPYSLPIVP-----GIVLRYTCSSYRLIGKRAIFGISENOVHAIW 66
Db      1/3 LQCEVVECDALNPVNVVTCPSLPSWNTTCACEKESFELIGPEHLQCTSS-----SSW 228
QY      57 KRAPPCISV-----NRTISCDSPLVHGIPMNGSKAPRHSQSVTFYTKANFTMK 107
Db      229 DKKKPKPCRAVTCVVGHPQNGVWSC-----NHSSGEPAYKSGHFTCAEGFCLQ 278
QY      108 GSKTVMCKANEMWGPTALPVGES 130
Db      279 GPAQI3CIAQSGWTQQA-PVCKA 300
RESULT 43
APOE CANFA
ID      APOILCANFA          STANDARD:      PRT:      345 AA.
AC      P13703;
DT      01-FEB-1994 (Rel. 28, Created)
DI      01-FEB-1994 (Rel. 28, last sequence update)
DI      16-OCT-2001 (Rel. 40, last annotation update)
DE      Beta-2-glycoprotein I precursor (Apolipoprotein II) (Apo-II) (B23PI)
DE      (Beta(2)GPI).

```


100

101

102

103

104

105

106

107

108

109

110

111

112

113

114

115

116

117

118

119

120

121

122

123

124

125

126

127

128

129

130

131

132

133

134

135

136

137

138

139

140

141

142

143

144

145

146

147

148

149

150

151

152

153

154

155

156

157

158

159

160

161

162

163

164

165

166

167

168

169

170

171

172

173

174

175

176

177

178

179

180

181

182

183

184

185

186

187

188

189

190

191

192

193

194

195

196

197

198

199

200

Genature version 5.1.3

Copyright (C) 1993 - 2002 CompuGen Ltd.

OM protein protein search, using sw model

Run on: November 6, 2002, 14:06:00 : Search time 21.5799 Seconds
(without alignments)
1042.224 Million cell updates/sec

Title: OS 09-834-309-6

Percent score: 75%

Sequence: 1 E1SCDPPPEVKNAKPKFYSL ANEMWPTALPVCESUPLLE 135

Scoring tables: BLOSUM62

Gapop 10.0, Gapext 0.5

Searched: 562222 seqs, 172994929 residues

Total number of hits satisfying chosen parameters: 562222

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post processing: Minimum Match 0%

Maximum Match 100%

Listing first 100 summaries

Database: SPTFMBL100

- 1: sp.archaea*
- 2: sp.bacteria*
- 3: sp.fungi*
- 4: sp.human*
- 5: sp.invertebrate*
- 6: sp.mammal*
- 7: sp.mhc*
- 8: sp.orquid*
- 9: sp.phage*
- 10: sp.plant*
- 11: sp.todent*
- 12: sp.virus*
- 13: sp.vertebrate*
- 14: sp.unclassified*
- 15: sp.virus*
- 16: sp.bacterioph*
- 17: sp.archae*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No	Score	Query Match	Length	DB ID	Description
1	752	100.0	1032	11 Q99883	Q99883 mus musculus
2	749	99.6	679	11 Q99254	Q99254 mus musculus
3	499	66.4	1045	6 Q46545	Q46545 ovis aries
4	466.5	62.0	1042	4 Q14866	Q14866 homo sapien
5	466.5	62.0	1047	4 Q14112	Q14112 homo sapien
6	241.5	40.8	1911	4 Q29528	Q29528 papio hamad
7	224	29.8	533	11 Q08569	Q08569 cavia porce
8	222	29.5	469	11 Q91848	Q91848 mus musculus
9	221.5	29.5	2014	6 Q29540	Q29540 pan troglod
10	221.5	29.5	2039	4 Q16745	Q16745 homo sapien
11	221.5	29.5	2489	4 Q16744	Q16744 homo sapien
12	220.5	29.3	132	4 Q9H099	Q9H099 homo sapien
13	220.5	29.3	132	4 Q9H098	Q9H098 homo sapien
14	219.5	29.2	132	4 Q9H055	Q9H055 homo sapien
15	215.5	28.7	132	4 Q9H054	Q9H054 homo sapien
16	215	28.6	579	11 Q60736	Q60736 mus musculus

17	214	28.5	555	11 Q997A1	Q997A1 cavia porce
18	199.5	26.5	363	6 Q02839	Q02839 sus scrofa
19	198.5	26.4	300	11 Q980R3	Q980R3 mus musculus
20	198.5	26.4	365	11 Q88174	Q88174 mus musculus
21	197.5	26.3	263	12 Q89855	Q89855 variola vir
22	197.5	26.3	263	12 Q07033	Q07033 variola vir
23	195.5	26.0	263	12 Q89074	Q89074 variola vir
24	195.5	26.0	263	12 Q89061	Q89061 variola vir
25	193.5	25.7	259	12 Q87613	Q87613 capox viru
26	188.5	25.1	349	4 Q15429	Q15429 homo sapien
27	185.5	24.7	645	12 Q9W802	Q9W802 bacula mola
28	184.5	24.5	285	6 Q19121	Q19121 papio hamad
29	184.5	24.5	1001	4 Q9Y2E1	Q9Y2E1 homo sapien
30	183.5	24.4	661	6 Q29531	Q29531 pan troglod
31	182.5	24.3	315	6 Q28770	Q28770 papio cynoe
32	181.5	24.1	285	6 Q19126	Q19126 macaca fasc
33	181.5	24.1	285	6 Q19127	Q19127 macaca mola
34	181.5	24.1	369	6 P79138	P79138 corysopithe
35	178.5	23.7	559	4 Q9H0V2	Q9H0V2 homo sapien
36	177	23.5	451	13 Q9DEG0	Q9DEG0 gallus gall
37	175.5	23.3	222	6 Q19120	Q19120 adus trivi
38	175.5	23.3	222	6 Q19125	Q19125 salicaria sol
39	175.5	23.3	346	6 Q62834	Q62834 salicaria sol
40	174	23.1	355	11 Q920M4	Q920M4 rattus norv
41	173	23.0	3567	11 Q9ES77	Q9ES77 mus musculus
42	172.5	22.9	417	11 Q35520	Q35520 rattus norv
43	172.5	22.9	497	11 Q36132	Q36132 rattus norv
44	172.5	22.9	559	11 Q63135	Q63135 rattus norv
45	172	22.9	1236	11 Q91Y86	Q91Y86 rattus norv
46	171.5	22.8	550	12 P88903	P88903 kapusi's sa
47	170.5	22.7	222	6 Q19122	Q19122 salicaria q
48	170	22.6	395	12 Q912M6	Q912M6 macaca mola
49	168.5	22.4	222	6 Q19124	Q19124 salicaria ny
50	168.5	22.4	314	6 Q62835	Q62835 salicaria ny
51	168.5	22.4	377	6 Q62836	Q62836 salicaria ny
52	168.5	22.4	378	6 Q62837	Q62837 salicaria ny
53	168	22.3	808	11 Q614C4	Q614C4 mus musculus
54	167.5	22.3	222	6 Q1912F	Q1912F salicaria p
55	167.5	22.3	225	12 Q91M83	Q91M83 lumpy's sa
56	167.5	22.3	550	12 Q40912	Q40912 kapusi's sa
57	167	22.2	483	11 Q64735	Q64735 mus musculus
58	166.5	22.1	522	6 Q28769	Q28769 papio cynoe
59	164.5	21.9	267	11 P97939	P97939 cavia porce
60	164.5	21.9	267	12 Q9DBG3	Q9DBG3 variola vir
61	164.5	21.9	337	11 P97263	P97263 variola vir
62	164.5	21.9	339	11 P97261	P97261 cavia porce
63	164.5	21.9	379	11 P70105	P70105 cavia porce
64	164.5	21.9	1653	5 Q9V109	Q9V109 drosophila
65	163.5	21.7	522	6 Q28797	Q28797 pan troglod
66	162.5	21.6	300	12 Q998T9	Q998T9 myxoma viru
67	162	21.5	711	4 Q9N015	Q9N015 homo sapien
68	162	21.5	716	4 Q95917	Q95917 homo sapien
69	161	21.4	777	4 Q9N013	Q9N013 homo sapien
70	161	21.4	787	4 Q9N014	Q9N014 homo sapien
71	161	21.4	792	4 Q9Y336	Q9Y336 homo sapien
72	161	21.4	1024	4 Q9BYH1	Q9BYH1 homo sapien
73	157.5	20.9	669	6 Q28085	Q28085 bras taurus
74	157	20.9	268	12 Q9QAX5	Q9QAX5 yata monkey
75	156.5	20.8	251	4 Q9ES25	Q9ES25 homo sapien
76	155.5	20.7	222	6 Q19123	Q19123 salicaria
77	155	20.6	3564	11 Q92313	Q92313 mus musculus
78	153.5	20.4	754	13 Q98315	Q98315 cypinus cu
79	153.5	20.4	754	4 Q9H86	Q9H86 homo sapien
80	153	20.3	569	4 Q9H866	Q9H866 homo sapien
81	151	20.1	923	11 Q91X64	Q91X64 mus musculus
82	150	19.9	340	11 Q62479	Q62479 mus musculus
83	150	19.9	359	13 Q73756	Q73756 trachydano
84	149.5	19.9	194	11 Q9QF08	Q9QF08 rattus norv
85	149	19.8	390	11 Q92130	Q92130 mus musculus
86	149	19.8	657	4 Q14006	Q14006 homo sapien
87	149	19.8	809	4 Q9H82	Q9H82 homo sapien
88	149	19.8	840	4 Q9H045	Q9H045 homo sapien
89	149	19.8	853	4 Q9H047	Q9H047 homo sapien


```

OX NCBI_TaxID=9548;
RN [1]
RP SEQUENCE FROM N.A.
EX MEDLINE=94292799; PubMed 8021505;
RA Birmingham D.J., Shen X.F., Bourcade D., Nickells M.W., Atkinson J.P.;
RT "Primary sequence of an alternatively spliced form of CRL1, Candidate
RT for the 75,000 M(r) complement receptor expressed on chimpanzee
RT erythrocytes."
RL J. Immunol. 153:691-700(1994).
DR EMBL: L24920; AAA51438.1;
DR HSSP: P08603; IHFI
DR InterPro: IPR001424; SOD_CU_ZN.
DR InterPro: IPR000436; Sushi_SCR_CCP.
DR Pfam: PF00084; sushi; 30.
DR SMART: SM00042; CCP; 30.
DR PROSITE: PS00143; CARBOXYPEPT_ZN_2; UNKNOWN_2.
DR PROSITE: PS00087; SOD_CU_ZN_1; UNKNOWN_1.
KW Receptor.
FT NON_TER 2014 2014
SQ SEQUENCE 2014 AA; 221281 MW; 6D6C3A74D81F1DB9 CRC64.

Query Match 29.58; Score 221.5; DB 6; Length 2014.
Best local Similarity 36.88; Pred. No. 4.9e-16;
Matches 50; Conservative 17; Mismatches 54; Indels 15; Gaps 5;

QY 2 ISCDPPEVKNAKPPYSS LPVETVLPYTCSPS-----YRLIGEKAIKICISNQV 52
DB 1492 ISCEPPPTISNG--DFYSNNKTSFNGTIVTYQCHTGHKQIFELVGRSIVCTSKDDQ 1549
QY 53 HATWKAADPTCESVKNKFTISCDPIVPGCFMKNKSKAPFHDSDVTFCKANFTMKSKIV 112
DB 1596 VGVWSSPPRCISTNK---CTAPEVENAIRVPGNRSEFSLTEIVRRCQPSFVWVSHIV 1606
QY 113 WQANEMMCPITALPVC 128
DB 1607 QCQTNGRWCP-KLPHC 1621

RESULT 10
Q16745 PRELIMINARY; PRT; 2039 AA.
AC Q16745;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE COMPLEMENT RECEPTOR 1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
EX MEDLINE=94065175; PubMed=8245463;
RA VIK D.P., Wong W.W.;
RT "Structure of the gene for the F allele of complement receptor type 1
RT and sequence of the coding region unique to the S allele."
RL J. Immunol. 151:6214-6224(1993).
RN [2]
RP SEQUENCE FROM N.A.
RA VIK D.P., Wong W.W.;
RL Submitted (JUN-1993) to the EMBL/GenBank/DBJ databases.
DR EMBL: L17418; AAB60694.1; JOINED.
DR EMBL: L17390; AAB60694.1; JOINED.
DR EMBL: L17499; AAB60694.1; JOINED.
DR EMBL: L17409; AAB60694.1; JOINED.
DR EMBL: L17419; AAB60694.1; JOINED.
DR EMBL: L17420; AAB60694.1; JOINED.
DR EMBL: L17421; AAB60694.1; JOINED.
DR EMBL: L17422; AAB60694.1; JOINED.
DR EMBL: L17423; AAB60694.1; JOINED.
DR EMBL: L17391; AAB60694.1; JOINED.

```

```

DR EMBL: L17392; AAB60694.1; JOINED.
DR EMBL: L17393; AAB60694.1; JOINED.
DR EMBL: L17394; AAB60694.1; JOINED.
DR EMBL: L17395; AAB60694.1; JOINED.
DR EMBL: L17396; AAB60694.1; JOINED.
DR EMBL: L17397; AAB60694.1; JOINED.
DR EMBL: L17398; AAB60694.1; JOINED.
DR EMBL: L17400; AAB60694.1; JOINED.
DR EMBL: L17401; AAB60694.1; JOINED.
DR EMBL: L17402; AAB60694.1; JOINED.
DR EMBL: L17403; AAB60694.1; JOINED.
DR EMBL: L17404; AAB60694.1; JOINED.
DR EMBL: L17405; AAB60694.1; JOINED.
DR EMBL: L17406; AAB60694.1; JOINED.
DR EMBL: L17407; AAB60694.1; JOINED.
DR EMBL: L17408; AAB60694.1; JOINED.
DR EMBL: L17410; AAB60694.1; JOINED.
DR EMBL: L17411; AAB60694.1; JOINED.
DR EMBL: L17412; AAB60694.1; JOINED.
DR EMBL: L17413; AAB60694.1; JOINED.
DR EMBL: L17414; AAB60694.1; JOINED.
DR EMBL: L17415; AAB60694.1; JOINED.
DR EMBL: L17416; AAB60694.1; JOINED.
DR HSSP: P08603; IHFI.
DR InterPro: IPR001424; SOD_CU_ZN.
DR InterPro: IPR000436; Sushi_SCR_CCP.
DR Pfam: PF00084; sushi; 30.
DR SMART: SM00042; CCP; 30.
DR PROSITE: PS00133; CARBOXYPEPT_ZN_2; UNKNOWN_2.
DR PROSITE: PS00087; SOD_CU_ZN_1; UNKNOWN_1.
KW Receptor.
SQ SEQUENCE 2039 AA; 223633 MW; B2FE311C63.6635 CRC64;

Query Match 29.58; Score 22.5; DB 4; Length 2039.
Best local Similarity 36.88; Pred. No. 4.9e-16;
Matches 50; Conservative 17; Mismatches 54; Indels 15; Gaps 5;

QY 2 ISCDPPEVKNAKPPYSS--LPVETVLPYTCSPS---YRLIGEKAIKICISNQV 52
DB 1517 ISCEPPPTISNG--DFYSNNKTSFNGTIVTYQCHTGHKQIFELVGRSIVCTSKDDQ 1574
QY 53 HATWKAADPTCESVKNKFTISCDPIVPGCFMKNKSKAPFHDSDVTFCKANFTMKSKIV 112
DB 1575 VGVWSSPPRCISTNK---CTAPEVENAIRVPGNRSEFSLTEIVRRCQPSFVWVSHIV 1631
QY 113 WQANEMMCPITALPVC 128
DB 1632 QCQTNGRWCP-KLPHC 1646

RESULT 11
Q16744 PRELIMINARY; PRT; 2419 AA.
AC Q16744;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE COMPLEMENT RECEPTOR 1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
EX MEDLINE=94065175; PubMed=8245463;
RA VIK D.P., Wong W.W.;
RT "Structure of the gene for the F allele of complement receptor type 1
RT and sequence of the coding region unique to the S allele."
RL J. Immunol. 151:6214-6224(1993).
RN [2]
RP SEQUENCE FROM N.A.

```



```

DB 282 SANNSEWEP-SIPKC 294
II I I I I I
RESULT 20
ID 088174 PRELIMINARY; PRT; 365 AA.
AC 088174;
DT 01 NOV 1998 (TRENBLREL, 08, Created)
DT 01 NOV 1998 (TRENBLREL, 08, Last sequence update)
DT 01 JUN 2001 (TRENBLREL, 17, Last annotation update)
DE MEMBRANE COFACTOR PROTEIN (CD46).
GN MFP.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Eumetazoa; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID:10090;
RN 111
RP SEQUENCE FROM N.A.
RA Itoji M., Kitamura M., Soya I.;
PT "Molecular cloning of a murine homologue of membrane cofactor protein
(CD46): preferential expression in testicular germ cells.";
PL Thesis (1997).
RL Osaka Medical Center for Cancer and Cardiovascular Diseases, Japan.
RN 121
RP SEQUENCE FROM N.A.
RA Masumoto M., Matsumiya K., Okuyama A., Nishimura Y., Okabe M.,
Soya I.;
PT "Molecular cloning of a murine homologue of membrane cofactor protein
(CD46): preferential expression in testicular germ cells.";
RL Biochem. J. 340:163-168(1998).
RN 141
RP SEQUENCE FROM N.A.
RA STRAIN-C57BL/6; TISSUE-TESTIS;
RX MEDLINE-94016109; PubMed-9799332;
RA Miwa T., Nonaka M., Okada N., Wakana S., Shiroishi T., Okada H.;
PT "Molecular cloning of rat and mouse membrane cofactor protein (MCP,
CD46): preferential expression in testis and close linkage between the
mouse Mcp and Cr2 genes on distal chromosome 1.";
PL Immunogenetics 48:463-471(1998).
RN 141
RP SEQUENCE FROM N.A.
RA STRAIN-C57BL/6J; TISSUE-TESTIS;
RX MEDLINE-21085660; PubMed-11217061;
RA Kawai J., Shinadawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
Aizawa K., Iizawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,
Saito T., Okazaki Y., Gajdardzic I., Aono H., Kasuoka T., Saito K.,
Kadoya K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
Fleischmann W., Gausterland T., Gissi C., King B., Kuchiwa H.,
Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,
Schriml I.M., Stambli F., Suzuki K., Tomita M., Wagner L., Washio T.,
Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Harsh G.,
Blake J., Boffelli D., Bolunda N., Carninci P., de Bonaldo M.P.,
Brownstein M.J., Bull C., Fletcher C., Fujita M., Gariboldi M.,
Gustafson S., Hill D., Haimann M., Hume D.A., Kamiya M., Lee N.H.,
Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombauts P.,
Nordone P., Rind H., Ringwald M., Rodriguez L., Sakamoto N.,
Sasaki H., Sato K., Schenck C., Soya T., Shibata Y., Storch K.-F.,
Suzuki H., Toyooka K., Waki K., Weier G., Whitaker S., Wilming L.,
Wyshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohsaki S.,
Hayashizaki Y.;
PT "Functional annotation of a full length mouse cDNA collection.";
RL Nature 409:685-690(2001).
DR EMBL: AB001566; HAA11859.1;
DR EMBL: AB010919; HAA34810.1;
DR EMBL: AK006642; BAB24682.1;
DR BSSP: pl0998; LVVD.
DR MGI: 1204290; Mcp.
DR InterPro: IPR000436; Sushi_SCR_CCP.
DR Pfam: PF00084; sushi; 4.

```

```

DR SMART: SM00042; CCP; 4.
KW Membrane, Signal.
SQ SEQUENCE 365 AA; 40881 MW; 84AA0A63E165C929 CRC64;
Query Match
Best Local Similarity 31.38; Score 196.5; DB 11; Length 465;
Matches 42; Conservative 22; Mismatches 55; Indels 15; Gaps 5;
QY 1 ELSQDIPPFVANKPKYYSILPIVPGTGLRYTPS - YPIISKALPTISENVVIA 4
DB 170 KIYCLIPPKIKNGTHITLDINVKYHEAVSYSCPTPIPKESILVHSMIPW3 - IIN 225
QY 35 IWKALPTICESVKNKIIISDPIVFGGFMNASKKAPFKHGSVIFGKANIIMAGSKIVW 114
DB 226 IWSNSPPK-----KVKKPFNVLQNCRLKAKGLIPYKQIVMEELGHIYMGSSMWIC 281
QY 115 QANEMMGHTALPVC 128
DB 282 SANNSEWEP-SIPKC 294
RESULT 21
Q89859 PRELIMINARY; PRT; 253 AA.
AC Q89859;
DT 01 NOV 1996 (TRENBLREL, 01, Created)
DT 01 NOV 1996 (TRENBLREL, 01, Last sequence update)
DT 01 JUN 2000 (TRENBLREL, 17, Last annotation update)
DE HOMOLOG OF VACCINIA VIRUS CDS CAL.
GN D151.
OS Variola virus.
OC Viruses; dsDNA viruses, no RNA stage; Poxviridae; Chordopoxvirinae;
OC Orthopoxvirus.
OX NCBI_TaxID:10255;
RN 111
RP SEQUENCE FROM N.A.
RA STRAIN-HANGLADESH-1975;
RX MEDLINE-94088747; PubMed-8264798;
RA Masung R.P., Esposito J.J., Liu L.I., Qi J., Utterback L.K.,
Knight J.C., Aubin L., Yuran T.E., Parsons T.M., Loparev V.N.;
PT "Potential virulence determinants in terminal regions of variola
smallpox virus genome";
RL Nature 366 748-751(1993).
RN 121
RP SEQUENCE FROM N.A.
RA STRAIN-SOMALIA-1977;
RA Masung R.P., Loparev V.N., Knight J.C., Chichikov V.E., Parsons T.M.,
Totmenin A.V., Shchelkunov S.N., Esposito J.J.;
RX Submitted (JUL-1995) to the EMBL/GenBank/DBJ databases.
DR EMBL: 122579; AAA60760.1;
DR EMBL: U18340; AAA69423.1;
DR BSSP: pl0948; LVVD.
DR InterPro: IPR000436; Sushi_SCR_CCP.
DR Pfam: PF00184; sushi; 4.
DR SMART: SM01032; CCP; 4.
SQ SEQUENCE 253 AA; 28816 MW; 04610C303027C4220 CRC64;
Query Match
Best Local Similarity 29.98; Score 197.5; DB 12; Length 264;
Matches 38; Conservative 20; Mismatches 58; Indels 11; Gaps 4;
QY 2 ISCDIPPFVANKPKYYSILPIVPGTGLRYTPS - YPIISKALPTISENVVIAIKRPP 61
DB 146 VKCQLPPIISNGRINGYNDFTDGSVWYTCNSYSLSINSGVLQSG-----GWSN 199
QY 62 ICESVKNKIIISDPIVFGGFMNASKKAPFKHGSVIFGKANIIMAGSKIVW 121
DB 200 TCG-----IVKCPHPPIILNGYISSGPKRSYVN NVDFTKYGKYLSSSSSSSTSGNHW 255
QY 122 FTALPVC 128
DB 256 P-ELPKC 261

```

[illegible]


```

Best Local Similarity 29.9%, Pred. No. 4, 1e-14;
Matches 48; Conservative 20; Mismatches 58; Indels 11; Gaps 4;

QY 2 ISDPPPEVKARKYYSLSPIVPGTIVLRYTCSYRLLIGKAIKPCISENVHATWCKAPP 61
DB 146 VKCQSPPSISNCRHNCYMDFTYSSVYTSNSYSLSIGNSCVLSC-----QFWSN-PP 199
QY 62 IFESVKNKITSDDPIVPGSGFMNKSAPFERHGSVTFETKANFTMKSGKIVWQGANIMWG 121
DB 200 TQY---IVKCHPILINWYSSSPKESYSYNQNVET-KWYKLSGSSSPSPINTWQ 255
QY 122 PTALPVC 128
DB 256 P ELPKC 261

RESULT 25
ID 187616 PRELIMINARY: PRT: 259 AA
AC 187616;
DT 01 MAY 1997 (Tremblrel, 64, Created)
DT 01 MAY 1997 (Tremblrel, 64, Last sequence update)
DT 01 JUN 2001 (Tremblrel, 17, Last annotation update)
DE 41KBP FRAGMENT FROM LEFT END OF GENOME.
GN 187616 OR C17L.
OS Coxsack virus (CPV).
OC Viruses; dsRNA viruses, no RNA stage; Poxviridae; Chordopoxvirinae;
OC Orthopoxvirus.
OX NCBI TaxID=10243;
RN 1;
RC SEQUENCE FROM N.A.
RE STRAIN=GR1 90;
RX MEDLINE=97068532; PubMed=8964248;
RA Saitonov P.F., Petrov N.A., Riazankina O.I., Totmenin A.V.,
RA Shebekunov S.N., Sandakchiev L.S.;
RT "Genes of a circle of hosts for the coxpos virus.",
RL Dokl. Akad. Nauk 349:829-833(1996);
RN 12;
RP SEQUENCE FROM N.A.
RE STRAIN=GR1 90;
RX MEDLINE=9829462; PubMed 9568042;
RA Shebekunov S.N., Saitonov P.F., Totmenin A.V., Petrov N.A.,
RA Ryzankina O.I., Gutorov V.V., Kotwal G.J.;
RT "The genomic sequence analysis of the left and right species-specific
RT terminal region of a coxpos virus strain reveals unique sequences and
RT a cluster of intact ORFs for immunomodulatory and host range
RT proteins."
RL Virology 243:432-440(1998)
DR EMBL: X94455; CAA64102.1;
DR EMBL: Y11842; CAA72567.1;
DR HSSP: P10998; 1VVD.
DR InterPro: IPR000436; Sushi_SCR_CCP.
DR Pfam: PF00084; sushi; 4.
DR SMART: SM00042; CCP; 4.
SQ SEQUENCE 259 AA; 28193 MW; 901AAEF6893B859A CRC64;

Query Match 25.7%; Score 193.5; DB 12; Length 259;
Best Local Similarity 29.9%, Pred. No. 6, 1e-14;
Matches 47; Conservative 19; Mismatches 60; Indels 11; Gaps 4;

QY 2 ISDPPPEVKARKYYSLSPIVPGTIVLRYTCSYRLLIGKAIKPCISENVHATWCKAPP 61
DB 142 VKCQSPPSISNCRHNCYMDFTYSSVYTSNSYSLSIGNSCVLSC-----QFWSN-PP 195
QY 62 IFESVKNKITSDDPIVPGSGFMNKSAPFERHGSVTFETKANFTMKSGKIVWQGANIMWG 121
DB 196 TQY---IVKCHPILINWYSSSPKESYSYNQNVET-KWYKLSGSSSPSPINTWQ 251
QY 122 PTALPVC 128
DB 252 P ELPKC 257

RESULT 26
ID 15429 PRELIMINARY: PRT: 649 AA
AC 15429;
DT 01 NOV-1996 (Tremblrel, 01, Created)
DT 01 NOV-1996 (Tremblrel, 01, Last sequence update)
DT 01 JUN-2001 (Tremblrel, 17, Last annotation update)
DE C146.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI TaxID=9606;
RN 1;
RC SEQUENCE FROM N.A.
RE TISSUE=TESTIS;
RA Baga T., Suzuki Y., Nagasawa S., Suya C.;
RT "cDNA cloning and characterization of human sperm C146."
RE Submitted (MAR-1996) to the EMBL/GenBank/DDBJ databases.
DR EMBL: D84105; BAA1224.1;
DR HSSP: P10998; 1VVD.
DR InterPro: IPR000436; Sushi_SCR_CCP.
DR Pfam: PF00084; sushi; 4.
DR SMART: SM00042; CCP; 4.
SQ SEQUENCE 349 AA; 39325 MW; REF EDA40D4C318E CRC64;

Query Match 25.1%; Score 183.5; DB 4; Length 449;
Best Local Similarity 30.6%; Pred. No. 3, 5e-14;
Matches 41; Conservative 23; Mismatches 55; Indels 1%; Gaps 6;

QY 1 ELSDDPIEVKNARKYYSLSPIVPG-IVLYVYSSS YRLEDKAIFGIFSNVHA 54
DB 159 KVLCTPKKIKGKHTESEVEFYLIAYLSJPAFCQPFESLDGSLVCGDLSV 215
QY 55 TWDKA-IPICESVKNKITSDDPIVPGSGFMNKSAPFERHGSVTFETKANFTMKSGKIVW 114
DB 216 -WSRAP-C-----NVVKCRHPVLENGVINGEJANFYKATVMPFEDKGYLDSHIV 270
QY 15 QANEMWGITALPVC 128
DB 271 DSNSTWDE-FVPC 283

RESULT 27
ID 99R02 PRELIMINARY: PRT: 545 AA.
AC 99R02;
DT 01 NOV-1994 (Tremblrel, 12, Created)
DT 01 NOV-1994 (Tremblrel, 12, Last sequence update)
DT 01 DEC-2001 (Tremblrel, 19, Last annotation update)
DE COMPLEMENT BINDING PROTEIN.
OS Macaca mulatta rhadinovirus 17577.
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Gammaherpesvirinae; Rhadinovirus.
OX NCBI TaxID=8534;
RN 1;
RC SEQUENCE FROM N.A.
RE MEDLINE=99174001; PubMed=10074154;
RA Seales R.P., Berquam E.P., Axthelm M.K., Wang S.W.;
RT "Sequence and genomic analysis of herpes macaque rhadinovirus with
RT similarity to Kaposi's sarcoma-associated Herpesvirus/human
RT herpesvirus 8."
RL J. Virol. 73 3040-3053(1999).
RN 12;
RP SEQUENCE FROM N.A.
RA Seales R.P., Berquam E.P., Axthelm M.K., Wang S.W.;
RE Submitted (JUN-2000) to the EMBL/GenBank/DDBJ databases.
DR EMBL: AF084501; AAD21332.1;
DR HSSP: P10998; 1VVD.
DR InterPro: IPR000436; Sushi_SCR_CCP.
DR Pfam: PF00084; sushi; 8.
DR SMART: SM00042; CCP; 8.
SQ SEQUENCE 645 AA; 71526 MW; 93B81F-5A8F61F82 CRC64;

```



```

DR SMART: SM00042; CCP; 9.
DR PROSITE: PS00087; SMO_CU_ZN_1; UNKNOWN_1.
KW Signal: Receptor.
FT NON_TER 1 1
FT SIGNAL 41 16 POTENTIAL.
FT CHAIN 17 661 COMPLEMENT RECEPTOR 1.
FT NON_TER 661 661
SQ SEQUENCE 661 AA; 72966 MW; 47009154266AD3P0 CRC64;

Query Match
Best Local Similarity 24.4%; Score 183.5; DB 6; Length 661;
Matches 45; Conservative 25; Mismatches 48; Indels 21; Gaps 7.

QY 3 SCDPPPEVKNARK PYSLPIVPGIVLRYTCSYSYRLIGKKAIFCLISENOVHATWKP 60
DB 3 111111 111111 111111 111111 111111 111111 111111 111111
469 A-PHPKPIUNGIDIGHSVL YLPQMTISYLCDPGYLLVKGKGFECTQ-----GINSQLD 523
QY 61 PICESVNVKLSQSDPIVNGFMNKGKSG-----KAPFRRGDSVFTTCRANKFMKSKIVWCQ 115
DB 111111 111111 111111 111111 111111 111111 111111 111111
524 HYCKEVN-----CSFPL-----FMNGISKLEMKKVVHYQDYVTKCRDGYTLKGPSWQOC 575
QY 116 ANEMMGPTALPVCS 140
DB 111111 111111 111111 111111 111111 111111 111111 111111
576 ADNRWUP PLAKGIS 589

RESULT 31
Q28770 PRELIMINARY; PRT; 315 AA.
ID Q28770
AC Q28770
DI 01 NOV 1996 (TREMblrel. 01, Created)
DT 01 NOV 1996 (TREMblrel. 01, Last sequence update)
DE COMPLEMENT RECEPTOR (FRAGMENT).
OS Papio cynocephalus (Yellow baboon).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;
OC Cercopithecoidea; Papio.
NCBI_TaxID=9556;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE-9746679; PubMed-9224509;
RA Hsu F.C., Yorig R.E., Sarangi F., Marill A., Iorio C.,
RA Richardson C.D.;
RA "Artificial mutations and natural variations in the CD46 molecules
RT from human and monkey cells define regions important for measles virus
RT binding.";
RL J. Virol. 71:6144-6154(1997).
DR FMBI; U87921; AAH66821.1; -.
DR HSD; P10958; LVVD.
DR InterPro: IPR000436; Sushi_SCR_CYP.
DR SMART: PF00084; sushi; 4.
DR SMART: SM00042; CCP; 5.
FT NON_TER 415 415
SQ SEQUENCE 415 AA; 45172 MW; 49581144E752824F CRC64;

Query Match
Best Local Similarity 32.4%; Score 182.5; DB 6; Length 315;
Matches 45; Conservative 18; Mismatches 57; Indels 19; Gaps 5;

QY 1 EISCDDPEVKNARKPYSLPIVPGIVLRYTCSYSYRLIGKKAIFCLISENOVHATWKP 49
DB 3 111111 111111 111111 111111 111111 111111 111111 111111
122 KISGLPPTDNGDSKFFIY-----GSVYTRNLGSGGKKLFELGEPISYCTSN 177
QY 50 NVVHATWKPATPIESVNVKLSQSDPIVNGFMNKGKSGKAPFRRGDSVFTTCRANKFMKSG 109
DB 111111 111111 111111 111111 111111 111111 111111 111111
178 EDVUGLWGPALQCLIPNK---CTPPNVENGILVSVNRSLFSLNEVFRRCPQCFVWKP 234
QY 110 KTVWKQANEMMGPTALPV 128
DB 111111 111111 111111 111111 111111 111111 111111 111111
245 RAVQVQALNKKWRP ELKSC 252

```

```

RESULT 32
Q19126 PRELIMINARY; PRT; 245 AA.
ID Q19126
AC Q19126
DI 01 JAN 1998 (TREMblrel. 05, Created)
DT 01 JAN 1998 (TREMblrel. 05, Last sequence update)
DE MEMBRANE COFACTOR PROTEIN (FRAGMENT).
OS Macaca fascicularis (Crab eating macaque) (Cynomolgus monkey).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;
OC Cercopithecoidea; Macaca.
NCBI_TaxID=9541;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE-9746679; PubMed-9224509;
RA Hsu F.C., Yorig R.E., Sarangi F., Marill A., Iorio C.,
RA Richardson C.D.;
RA "Artificial mutations and natural variations in the CD46 molecules
RT from human and monkey cells define regions important for measles virus
RT binding.";
RL J. Virol. 71:6144-6154(1997).
DR FMBI; U87921; AAH66821.1; -.
DR HSD; P10958; LVVD.
DR InterPro: IPR000436; Sushi_SCR_CYP.
DR SMART: PF00084; sushi; 4.
DR SMART: SM00042; CCP; 4.
FT NON_TER 245 245
SQ SEQUENCE 245 AA; 31998 MW; 62910413ED54E1E4 CRC64;

Query Match
Best Local Similarity 29.1%; Score 181.5; DB 6; Length 285;
Matches 39; Conservative 25; Mismatches 55; Indels 15; Gaps 5;

QY 1 EISCDDPEVKNARKPYSLPIVPGIVLRYTCSYSYRLIGKKAIFCLISENOVHATWKP 54
DB 3 111111 111111 111111 111111 111111 111111 111111 111111
159 KILCTPPPIKNGKHTFSEVFEVYLLAVIYSCAPAGCPQPSLIGESMIYQCN NS 214
QY 35 TWDKAVPICESVNVKNTISGSDPIVPGIVLRYTCSYSYRLIGKKAIFCLISENOVHATWKP 114
DB 111111 111111 111111 111111 111111 111111 111111 111111
215 TWSHAAPPC---KVKCRFPVVENKQISGFGKKFYKATVMEFCDKGYVINGSKIVIC 270
QY 115 QANEMMGPTALPV 128
DB 3 111111 111111 111111 111111 111111 111111 111111 111111
271 ESNSTNDP-PVPRK 283

RESULT 33
Q19127 PRELIMINARY; PRT; 245 AA.
ID Q19127
AC Q19127
DI 01 JAN 1998 (TREMblrel. 05, Created)
DT 01 JAN 1998 (TREMblrel. 05, Last sequence update)
DE MEMBRANE COFACTOR PROTEIN (FRAGMENT).
OS Macaca mulatta (Rhesus macaque).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;
OC Cercopithecoidea; Macaca.
NCBI_TaxID=9544;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE-9746679; PubMed-9224509;
RA Hsu F.C., Yorig R.E., Sarangi F., Marill A., Iorio C.,
RA Richardson C.D.;
RA "Artificial mutations and natural variations in the CD46 molecules
RT from human and monkey cells define regions important for measles virus
RT binding.";
RL J. Virol. 71:6144-6154(1997).
DR FMBI; U87922; AAH66822.1; -.

```



```

OS Gallus gallus (Chicken);
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neornithae; Galliformes; Phasianidae; Phasianinae;
OC Gallus;
OX NCBI_TaxID=9041;
RN [1]
RP SEQUENCE FROM N.A.
RX MFI INF=95714; PubMed=112439;
RA Inoue N., Fukui A., Nomura M., Matsumoto M., Toyoshima K., Soga T.;
RT "A novel chicken membrane associated complement regulatory protein;
RI Molecular cloning and functional characterization.";
RE J. Immunol. 166:424-431(2001);
DR EMBL: AR045592; RAB16878.1;
DR HSSP: P10998; 1VVD;
DR InterPro: IPR000436; Sushi_SCR_CDP;
DR Pfam: PF00084; sushi; 6;
DR SMART: SM00042; CCP; 6;
SQ SEQUENCE 451 AA; 48723 MW; 41DE1219A502837 CRC64.

Query Match 23.5%; Score 177; DB 13; Length 451;
Best Local Similarity 28.4%; Prod. No. 9.5e-12;
Matches 38; Conservative 21; Mismatches 63; Indels 12; Gaps 4.

QY 2 ISCDPPPEVKNARKYYSLSIPVIGVTLVLTGSP-----SYRLGKAIPTISENVVHA 55
DB 148 VPKRPPSTANGRYT-FAANYVYQTTVTYSCTDVRTGKNSPSISFTVDENSNV 206
QY 56 WKAPPTIESVKNKTSISDPIVPGSEMKKAPRHCDSVTECKANFTMKGSKTYVW 114
DB 207 WSGPPPPQ-----KVVKCRFPVIVENGROTA:EGTSPYNAAMVPEYFTHIESH 267
QY 116 ANEMWGPATLPVCE 129
DB 263 ENGWTWP-SLTPCK 275

RESULT 37
ID 019120 PRELIMINARY; PRT; 222 AA.
AC 019120;
DT 01 JAN 1998 (TRENBLREL. 05, Created)
DT 01 JAN 1998 (TRENBLREL. 05, Last sequence update)
DT 01 JUN 2001 (TRENBLREL. 17, Last annotation update)
DE MEMBRANE COFACTOR PROTEIN (FRAGMENT).
GN CD46.
OS Aotus trivirgatus (Night monkey) (Douroucoulli);
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Platyrrhini; Cebidae; Aotinae; Aotus.
OX NCBI_TaxID=9505;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=97466679; PubMed=9224509;
RA Hsu E.C., Dorris K.E., Sarangi F., Marell A., Torio C.;
RT "Artificial mutations and natural variations in the cp46 molecules
RI from human and monkey cells define regions important for measles virus
binding.";
RE J. Virol. 71:6144-6154(1997);
DR EMBL: U87914; AAB66814.1;
DR HSSP: P10998; 1VVD;
DR InterPro: IPR000436; Sushi_SCR_CDP;
DR Pfam: PF00084; sushi; 3;
DR SMART: SM00042; CCP; 3;
DR NON_TER 222 222
SQ SEQUENCE 222 AA; 24469 MW; 192AC474F0D0BFD0D CRC64.

Query Match 23.4%; Score 175.5; DB 6; Length 222;
Best Local Similarity 28.4%; Prod. No. 6.2e-12;
Matches 38; Conservative 22; Mismatches 59; Indels 15; Gaps 5;

QY 1 EISCDPPPEVKNARKYYSLSIPVIGVTLVLTGSP-----YRLGKAIPTISENVVHA 54
DB 96 KILCKPPTIKNGKHTSEVIVFVYLLAVTYSC:PAKGPDPFSLVGSTIYV----RDSL 151
QY 55 TWKAPPTIESVKNKTSISDPIVPGSEMKKSKAPRHCDSVTECKANFTMKGSKTYVW 114
DB 152 GWSGDPPE-----KVVKCRFPVIVENGROTA:EGTSPYNAAMVPEYFTHIESH 267
QY 115 QANEMWGPATLPVCE 128
DB 208 NSNSTWDP-FVPEKC 220

RESULT 38
ID 062834 PRELIMINARY; PRT; 345 AA.
AC 062834;
DT 01 AUG 1998 (TRENBLREL. 07, Created)
DT 01 AUG 1998 (TRENBLREL. 07, Last sequence update)
DT 01 DEC 2001 (TRENBLREL. 19, Last annotation update)
DE MEMBRANE COFACTOR PROTEIN CD46.
GN CD46.
OS Sacalinus oedipus (Cotton-top tamarin);
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Platyrrhini; Callitrichidae; Saimiri.
OX NCBI_TaxID=9490;
RN [1]
RP SEQUENCE FROM N.A.
RX STEAIN-B95-6;
RX MFI INF=98161817; PubMed=9494106;

```

```

QY 55 TWKAPPTIESVKNKTSISDPIVPGSEMKKSKAPRHCDSVTECKANFTMKGSKTYVW 114
DB 152 GWSGDPPE-----KVVKCRFPVIVENGROTA:EGTSPYNAAMVPEYFTHIESH 267
QY 115 QANEMWGPATLPVCE 128
DB 208 NSNSTWDP-FVPEKC 220

RESULT 38
ID 019125 PRELIMINARY; PRT; 222 AA.
AC 019125;
DT 01 JAN 1998 (TRENBLREL. 05, Created)
DT 01 JAN 1998 (TRENBLREL. 05, Last sequence update)
DT 01 JUN 2001 (TRENBLREL. 17, Last annotation update)
DE MEMBRANE COFACTOR PROTEIN (FRAGMENT).
GN CD46.
OS Saimiri sciureus (Common squirrel monkey);
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Platyrrhini; Cebidae; Saimiri.
OX NCBI_TaxID=9521;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=97356479; PubMed=9224509;
RA Hsu E.C., Dorris K.E., Sarangi F., Marell A., Torio C.;
RT "Artificial mutations and natural variations in the CD46 molecules
RI from human and monkey cells define regions important for measles virus
binding.";
RE J. Virol. 71:6144-6154(1997);
DR EMBL: U87914; AAB66814.1;
DR HSSP: P10998; 1VVD;
DR InterPro: IPR000436; Sushi_SCR_CDP;
DR Pfam: PF00084; sushi; 3;
DR SMART: SM00042; CCP; 3;
DR NON_TER 222 222
SQ SEQUENCE 222 AA; 24517 MW; 75DE1D75L8B8378 CRC64.

Query Match 23.3%; Score 75.5; DB 6; Length 222;
Best Local Similarity 28.4%; Prod. No. 6.2e-12;
Matches 38; Conservative 22; Mismatches 59; Indels 15; Gaps 5;

QY 1 EISCDPPPEVKNARKYYSLSIPVIGVTLVLTGSP-----YRLGKAIPTISENVVHA 54
DB 96 KILCKPPTIKNGKHTSEVIVFVYLLAVTYSC:PAKGPDPFSLVGSTIYV----RDSL 151
QY 55 TWKAPPTIESVKNKTSISDPIVPGSEMKKSKAPRHCDSVTECKANFTMKGSKTYVW 114
DB 152 GWSGDPPE-----KVVKCRFPVIVENGROTA:EGTSPYNAAMVPEYFTHIESH 267
QY 115 QANEMWGPATLPVCE 128
DB 208 NSNSTWDP-FVPEKC 220

RESULT 39
ID 062834 PRELIMINARY; PRT; 345 AA.
AC 062834;
DT 01 AUG 1998 (TRENBLREL. 07, Created)
DT 01 AUG 1998 (TRENBLREL. 07, Last sequence update)
DT 01 DEC 2001 (TRENBLREL. 19, Last annotation update)
DE MEMBRANE COFACTOR PROTEIN CD46.
GN CD46.
OS Saimiri sciureus (Common squirrel monkey);
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Platyrrhini; Cebidae; Saimiri.
OX NCBI_TaxID=9490;
RN [1]
RP SEQUENCE FROM N.A.
RX STEAIN-B95-6;
RX MFI INF=98161817; PubMed=9494106;

```


RESULT 42

045520
AC 035520 PRELIMINARY; PRI: 417 AA.
01 JAN 1998 (TREMblrel. 01, Created)
01 JAN 1998 (TREMblrel. 01, Last sequence update)
01 DEC 2001 (TREMblrel. 16, Last annotation update)
DE 512 ANTIGEN (FRAGMENT).
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID-10116;
RN [1]
RF SEQUENCE FROM N.A.
RC TISSUE=LIVER;
RA Dohi N., Sakurada C., Nonaka M., Okada N., Okada H.;
RI "Novel isoform of rat complement regulatory protein, rat crry.";
RL Submitted (Nov 1994) to the EMBL/GenBank/DBJ databases.
DR EMBL: D42116; BAA22549.1;
DR HSSP: P10998; 1VVD.
DR InterPro: IPR000436; Sushi_SCR_CCP.
DR Pfam: PF00084; sushi; 6.
DR SMART: SM00042; CCP; 6.
FT NON TER 417 417.
SQ SEQUENCE 417 AA: 45951 MW: 042046609EB9BDB CRC64;

Query Match 22.9%; Score 172.5; DB 11; Length 417;
Best Local Similarity 30.2%; Pred. No. 2, Be-11;
Matches 42; Conservative 21; Mismatches 55; Indels 21; Gaps 5;
QY 2 ISCDPPPEVKNA-----RKPYSLPIVPGIVLRTCTSPS-----YRLGKKAIFPWISE 49
DB 159 IPCEIPSPSPNGDFSPNREDFH-----GMVVYQCNTIARGKKLENVCEPSIHCTSI 75
QY 50 NOVHATWIKAPICSVNKTISCSDPIVPGIIPVGGFMNKGSKAPRHDGVSITFTCKANFTMKGS 109
DB 76 DGGGVWMSGFPQCETELNK---CTPPHVENAVIVSKKSLFSLRQWVFPPGQCFPMKKGQD 132
QY 110 KIVWQANPMWGPALPVC 128
DB 133 SSVYGRSLNHWEP-QLPSC 288

RESULT 43

063612
AC 063612 PRELIMINARY; PRI: 497 AA.
01 NOV 1996 (TREMblrel. 01, Created)
01 NOV 1996 (TREMblrel. 01, Last sequence update)
01 JUN 2001 (TREMblrel. 17, Last annotation update)
DE 512 ANTIGEN (PRECURSOR).
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID-10116;
RN [1]
RF SEQUENCE FROM N.A.
RC TISSUE=LIVER;
RA Sakurada C., Seno H., Dohi N., Takizawa H., Nonaka M., Okada N.,
RI "Molecular cloning of the rat complement regulatory protein, 512 antigen.";
RL Biochem. Biophys. Res. Commun. 198;819-826(1994).
DR EMBL: D42114; BAA07698.1;
DR HSSP: P10998; 1VVD.
DR InterPro: IPR000436; Sushi_SCR_CCP.
DR Pfam: PF00084; sushi; 6.
DR SMART: SM00042; CCP; 6.
FT SIGNAL 1 46 POTENTIAL.
FT CHAIN 37 497 512 ANTIGEN.
SQ SEQUENCE 497 AA: 54786 MW: 177AC11EE0FAD1C CRC64;

Query Match 22.9%; Score 172.5; DB 11; Length 497;
Best Local Similarity 30.2%; Pred. No. 3,4e-11;
Matches 42; Conservative 21; Mismatches 55; Indels 21; Gaps 5;
QY 2 ISCDPPPEVKNA-----RKPYSLPIVPGIVLRTCTSPS-----YRLGKKAIFPWISE 49
DB 159 IPCEIPSPSPNGDFSPNREDFH-----GMVVYQCNTIARGKKLENVCEPSIHCTSI 214
QY 50 NOVHATWIKAPICSVNKTISCSDPIVPGIIPVGGFMNKGSKAPRHDGVSITFTCKANFTMKGS 109
DB 214 DGGGVWMSGFPQCETELNK---CTPPHVENAVIVSKKSLFSLRQWVFPPGQCFPMKKGQD 276
QY 110 KIVWQANPMWGPALPVC 128
DB 271 SSVYGRSLNHWEP-QLPSC 288

RESULT 44

063135
AC 063135 PRELIMINARY; PRI: 554 AA.
01 NOV-1996 (TREMblrel. 01, Created)
01 NOV-1996 (TREMblrel. 01, Last sequence update)
01 JUN-2001 (TREMblrel. 17, Last annotation update)
DE COMPLEMENT REGULATORY PROTEIN.
GN CRRY.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID-10116;
RN [1]
RF SEQUENCE FROM N.A.
RC STRAIN=SPRACD-DAWLEY;
RX MEDLINE-96006570. PubMed-7590969;
RA Guica R.J., Le C.F., Alexander J.J., Streed A.E., Moxley G. III;
RI "Molecular characterization of rat crry, widespread distribution of two alternative forms of crry mRNA.";
RL Immunogenetics 42:362-367(1995).
RN 2;
RF SEQUENCE FROM N.A.
RC TISSUE=LIVER;
RA Dohi N., Sakurada C., Nonaka M., Okada N., Okada H.;
RI Submitted (Oct 1997) to the EMBL/GenBank/DBJ databases.
DR EMB: 136532; AAA91821.1;
DR EMB: D42115; BAA22548.1;
DR HSSP: P10998; 1VVD.
DR InterPro: IPR000436; Sushi_SCR_CCP.
DR Pfam: PF00084; sushi; 7.
DR SMART: SM00042; CCP; 7.
SQ SEQUENCE 559 AA: 61680 MW: 29E10F6A210B96E CRC64;

Query Match 22.9%; Score 172.5; DB 11; Length 559;
Best Local Similarity 30.2%; Pred. No. 3,9e-11;
Matches 42; Conservative 21; Mismatches 55; Indels 21; Gaps 5;
QY 2 ISCDPPPEVKNA-----RKPYSLPIVPGIVLRTCTSPS-----YRLGKKAIFPWISE 49
DB 159 IPCEIPSPSPNGDFSPNREDFH-----GMVVYQCNTIARGKKLENVCEPSIHCTSI 214
QY 50 NOVHATWIKAPICSVNKTISCSDPIVPGIIPVGGFMNKGSKAPRHDGVSITFTCKANFTMKGS 109
DB 214 DGGGVWMSGFPQCETELNK---CTPPHVENAVIVSKKSLFSLRQWVFPPGQCFPMKKGQD 276
QY 110 KIVWQANPMWGPALPVC 128
DB 271 SSVYGRSLNHWEP-QLPSC 288
RESULT 45
Q91YB6
ID Q91YB6 PRELIMINARY; PRI: 1236 AA.
AC Q91YB6;

